

Tight bounds on the simultaneous estimation of incompatible parameters

Jasminder S. Sidhu,^{1,2,*} Yingkai Ouyang,^{1,†} Earl T. Campbell,¹ and Pieter Kok¹

¹*Department of Physics and Astronomy, The University of Sheffield, Sheffield, S3 7RH, UK*

²*SUPA Department of Physics, The University of Strathclyde, Glasgow, G4 0NG, UK*

(Dated: August 7, 2020)

The estimation of multiple parameters in quantum metrology is important for a vast array of applications in quantum information processing. However, the unattainability of fundamental precision bounds for incompatible observables has greatly diminished the applicability of estimation theory in many practical implementations. The Holevo Cramér-Rao bound (HCRB) provides the most fundamental, simultaneously attainable bound for multi-parameter estimation problems. A general closed form for the HCRB is not known given that it requires a complex optimisation over multiple variables. In this work, we develop an analytic approach to solving the HCRB for two parameters. Our analysis reveals the role of the HCRB and its interplay with alternative bounds in estimation theory. For more parameters, we generate a lower bound to the HCRB. Our work greatly reduces the complexity of determining the HCRB to solving a set of linear equations that even numerically permits a quadratic speedup over previous state-of-the-art approaches. We apply our results to compare the performance of different probe states in magnetic field sensing, and characterise the performance of state tomography on the codespace of noisy bosonic error-correcting codes. The sensitivity of state tomography on noisy binomial codestates can be improved by tuning two coding parameters that relate to the number of correctable phase and amplitude damping errors. Our work provides fundamental insights and makes significant progress towards the estimation of multiple incompatible observables.

I. INTRODUCTION

Physical quantities such as time, phase, and entanglement cannot be measured directly, but instead must be inferred through indirect measurements. An important category of such indirect measurements is parameter estimation. Quantum metrology describes the quantum mechanical framework that handles this estimation procedure. By recasting the problem as a statistical inference problem, parameter estimation can be associated with fundamental precision bounds. The key question in quantum metrology is what is the fundamental precision bound and how we can achieve it. Early applications of estimation theory focused on single parameter estimation such as phase measurements [1–3]. The ultimate precision bound for a single parameter is the quantum Cramér-Rao lower bound (QCRB), which was proved by Helstrom and Holevo [4–6]. Multi-parameter quantum metrology extends the single parameter case [7–9], and is of fundamental importance in understanding a variety of practical applications, such as Hamiltonian tomography [10], field sensing [11–13] and imaging [14–18], and distributed sensing [19–22]. A central problem is to determine the optimal measurement strategies that saturate the QCRB [23]. To achieve this, one must assume locally unbiased estimators [24], which is reasonable given large amounts of prior information [25, 26], and with many independent repetitions of the experiment [27]. Several reviews on the topic highlight recent progress in the field [28–30].

Each individual parameter we wish to estimate has an optimal measurement observable. However, when we wish to estimate two or more parameters simultaneously, the corresponding optimal observables may be incompatible. In this

case, we can not achieve the optimal precision for each parameter individually. In this case the QCRB matrix bound is generally not simultaneously saturable for all parameters [31–33]. This motivates the search for tighter bounds that can be realised for practical applications of multi-parameter estimation theory. The Holevo Cramér Rao bound (HCRB) encapsulates the difficulties associated with incompatible observables [6]. Operationally, it is the maximum of all existing lower bounds for the error of unbiased measurements [6]. It represents the best precision attainable with collective measurements on an asymptotically large number of identical copies of a quantum state [34–37].

Despite its importance, the HCRB has seen limited use in quantum metrology so far. There are several reasons for this. First, the HCRB is difficult to evaluate given that it is defined through a complex optimisation over a set of observables. Second, implementing collective measurements is generally a difficult task. Nevertheless, applications of the HCRB in metrological tasks do exist. Suzuki found closed form results for parameter estimation with qubits [38], and explored connections between different types of metrological bounds in the special case of two parameter estimation theory. For pure states [7] and displacement estimation with Gaussian states [6], it has been shown that the HCRB is attained by single-copy measurements. The HCRB was also used as a tool to define the precision of state estimation for finite dimensional quantum systems [37]. Bradshaw *et al.* calculated the HCRB for a joint parameter estimation of a displacement operation on a pure two-mode squeezed probe [39].

Arguably, the HCRB is most relevant in multi-parameter estimation. An increasing number of true multi-parameter estimation protocols has been explored [40–43], and therefore the need for general, attainable bounds on multi-parameter quantum estimation is urgent. Recently, Albarelli *et al.* have investigated the numerical tractability of calculating the HCRB for the simultaneous estimation of multiple parameters [44].

*Electronic address: jsmrdsidhu@gmail.com

†Electronic address: y.ouyang@sheffield.ac.uk

For finite-dimensional systems, they recast the evaluation of the HCRB as a semi-definite program, which is an optimisation problem that can be efficiently implemented. To date, no general analytic expression for the HCRB is known.

In this paper, we find that it is possible to recast the HCRB as a quadratic program with linear constraints, thereby providing tight bounds for multi-parameter estimation problems. We develop an analytical approach to solving the two-parameter HCRB, and provide expressions on when the analytical solution is tight. Our analytical solution for the optimal observables that can saturate the HCRB allows one to establish analytically the minimum penalty due to the incompatibility of the observables. Specifically, we generalise attainability constraints for simultaneous multi-parameter estimation problems where the commonly used Cramér-Rao bounds can not be saturated due to incompatibility. The analytic two-parameter HCRB can be considered a generalised quantum uncertainty relation [45]. For more than two parameters, our method does not provide tight bounds but still outperforms the QCRB.

A. Summary of results

The HCRB is defined as a constrained minimisation problem over measurement observables. By recasting the definition as a quadratic program with linear constraints, we find exact solutions to this minimisation and determine the optimal observables. Our method to solve this minimisation relies on the notion of duality in optimisation theory, where the primal problem is transformed to its dual problem. Through the duality gap, we are able to quantify the minimum penalty of estimating incompatible observables.

In this article, we introduce three new algorithms that derive bounds to the HCRB for two and more parameters. First we determine upper and lower bounds to the HCRB for two-parameter estimation problems that are not always tight. This leads to simple analytic expressions that are straightforwardly determined for probe states with full rank. The salient feature of this framework, from which the simplification is inherited, is that only the boundary values for the Lagrange dual variables are considered. Second, this method is extended to determine upper and lower bounds to the HCRB for more than two parameters. Finally, we return to the two-parameter HCRB to develop tight bounds. For this, we lift the constraint on the values for the Lagrange dual variables to explore the full generality permitted by our method. Our analysis for this shows that the HCRB is a general solution to a Sylvester equation in the measurement observables, and recovers the standard Lyapunov SLD QCRB solution when the weak commutativity criterion is violated. This algorithm can be implemented numerically using a Bartel-Stewart's algorithm for linear equation solvers, and offers a quadratic speedup in runtime over state-of-the-art semi-definite programming approaches.

Table I provides a high level summary of these algorithms, along with any assumptions made. Our results provide a significant extension of the capabilities of previous approaches, and clarifies the role of the HCRB in the estimation of incompatible observables.

B. Outline of paper

We begin in section II by providing an overview of multi-parameter quantum estimation. In section III, we introduce the four new algorithms for analytic and numerical results to the HCRB for two parameters and arbitrary number of parameters. We detail connections between alternative precision bounds and significantly extend the capabilities of previous approaches in the literature. Section IV discusses applications of our results to magnetometry and explores how bosonic quantum codes can bestow resilience of parameter estimates against noise beyond practical control. These applications demonstrate the strengths of our results and extend deep connections between quantum metrology and quantum error correcting codes. Finally, conclusions and interesting extensions to our results are provided in section V.

II. MULTI-PARAMETER QUANTUM ESTIMATION

Quantum estimation theory provides fundamental bounds to the estimation precision of physical parameters and the optimal measurements that saturate these limits [23]. We are interested in estimating multiple parameters simultaneously. The prototypical scheme requires that the vector of parameters $\theta = (\theta_1, \dots, \theta_d)^\top \in \mathbb{R}^d$ be imprinted on a quantum state $\rho(\theta)$. Denoting \mathbb{H}_D as the set of all Hermitian matrices in the Hilbert space of dimension D , we can see that $\rho(\theta)$ is a positive semidefinite matrix in \mathbb{H}_D with unit trace. We define measurement operators via a positive operator valued measure (POVM)

$$\Pi = \left\{ \Pi_\omega \geq 0, \omega \in \Omega \mid \sum_{\omega \in \Omega} \Pi_\omega = \mathbb{1}_D \right\}, \quad (1)$$

where $\mathbb{1}_D$ denotes the identity operator, and Ω is the set of measurement outcomes. The outcomes of such a measurement can be used in a function called the estimator $\check{\theta}$, which gives an estimate of the parameters. A general estimation scheme requires access to multiple identical copies of the quantum probe state. A separable measurement can be individually applied to each copy of the state to obtain estimates of each parameter separately, whereas a collective measurement can be applied jointly on all copies of the state to acquire a simultaneous estimate of all parameters. The ultimate precision bound is the one that is asymptotically achieved by a sequence of the best collective measurements as the number of copies tends to infinity [9, 34–36, 46–50].

The performance of the estimator $\check{\theta}$ under any measurement can be quantified in terms of its mean square estimate matrix (MSEM)

$$\Sigma_\theta(\Pi, \check{\theta}) = \sum_{\omega \in \Omega^N} p(\omega|\theta) (\check{\theta}(\omega) - \theta) (\check{\theta}(\omega) - \theta)^\top, \quad (2)$$

where the probability of measurement outcomes is provided by Born's rule $p(\omega|\theta) = \text{Tr}[\rho(\theta)\Pi_\omega]$, and N is the number of

Nature of HCRB bound	Assumptions	Algorithm details
Algorithm 1: analytic two-parameter bound.	Full rank ρ , linearly independent ρ_1 and ρ_2 , and analytic form for Q -matrix.	Provides upper \mathcal{U} and lower \mathcal{L} bounds, need not be tight.
Algorithms 2 and 3: hybrid multi-parameter bound.	Full rank ρ and analytic form for Q -matrix.	Provides upper \mathcal{U} and lower \mathcal{L} bounds, need not be tight.
Algorithm 4: hybrid two-parameter bound.	Full rank ρ , spectral decomposition of ρ . Full rankness of intermediate. Q -matrix is full rank and takes at most τ time to compute.	Analytic bounds for $u \in [0, 1]$. Tight bounds certifiably attained by numerically varying u to maximise \mathcal{L}_u . Computes in $O(\text{polylog}(1/\epsilon)\tau D^0)$.
Numerical two-parameter bound using Eq. (34).	Full rank ρ .	Computes in $O(\text{polylog}(1/\epsilon)D^{2.376})$ time using Bartel-Stewart's algorithm, or $O(\text{polylog}(1/\epsilon)D^3)$ time using Gaussian elimination.
SDP numerical algorithm.	Arbitrary ρ , ρ_1 , and ρ_2 .	Computes in $O(\text{polylog}(1/\epsilon)D^{2 \times 2.376})$ time, or $O(\text{polylog}(1/\epsilon)D^6)$ time using Gaussian elimination.

TABLE I: Algorithms and bounds to the HCRB for two and more parameters. Bounds are analytic, numerical, or a hybrid of analytical and numerical, as indicated in the first column. There is a trade-off between the assumptions taken for each algorithm and its complexity. Here D is the dimension of the probe state, and ϵ a measure of how close the HCRB bound is to optimal. The final row provides comparative details for the SDP approach in Ref [44].

independently repeated measurements. The set of estimators are said to be locally unbiased if for all $\omega \in \Omega$

$$\sum_{\omega \in \Omega^N} (\check{\theta}_j(\omega) - \theta_j) p(\omega|\theta) = 0, \quad \sum_{\omega \in \Omega^N} \check{\theta}_j(\omega) \partial_k p(\omega|\theta) = \delta_{jk}, \quad (3)$$

where $\partial_k \equiv \partial/\partial\theta_k$. Under these conditions, the MSEM is equivalent to the covariance matrix of parameter estimates, and is lower bounded through generalisations of the Cramér-Rao bound from classical statistics

$$\Sigma_\theta(\Pi, \check{\theta}) \geq \mathcal{F}(\rho(\theta), \Pi)^{-1}, \quad (4)$$

where \mathcal{F} is the classical Fisher information matrix [5, 6]. The Fisher information characterises the MSEM for the best classical data manipulation given a measurement strategy in the asymptotic limit [27]. A well known quantum generalisation includes the symmetric logarithmic derivative (SLD), $L_j \in \mathbb{H}_N$, which is implicitly defined through $2\partial_j \rho = \{L_j, \rho\}$ and generates the real symmetric quantum Fisher information (QFIM) $\mathcal{I}_{jk}^S = \text{Re}[\text{Tr}[\rho L_j L_k]]$ [4, 51]. This is referred to as the SLD QFIM. Notice that for ease of notation, we drop the explicit dependence of the state on the vector of parameters θ . Similarly, the right logarithmic derivative (RLD), $R_j \in \mathbb{H}_N$, defined through $\partial_j \rho = \rho R_j$, induces the complex Hermitian RLD QFIM $\mathcal{I}_{jk}^R = \text{Tr}[\rho R_j R_k]$ [52, 53]. The matrices \mathcal{I}^S and \mathcal{I}^R generate different lower bounds with a corresponding scalar SLD QCRB and RLD QCRB, defined via

$$C_S(\rho, \theta, W) = \text{Tr}[W[\mathcal{I}^S]^{-1}], \quad (5)$$

$$C_R(\rho, \theta, W) = \text{Tr}[W \text{Re}[\mathcal{I}^R]^{-1}] + \|\sqrt{W} \text{Im}[\mathcal{I}^R]^{-1} \sqrt{W}\|_1. \quad (6)$$

Here the weight matrix W is a size D positive definite square matrix that satisfies $x^\top W x > 0$ for all vectors $x \neq 0$, and $\|\cdot\|_1$

denotes the trace norm of a matrix [54]. Throughout, $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ of a matrix denotes taking the real and imaginary part of each matrix element. Nagaoka investigated in detail the relationship between these bounds [55]. The central problem in quantum estimation theory is the minimisation of these scalar bounds over the family of probability of distributions defined by quantum measurements.

Helstrom demonstrated that the SLD QCRB is tight for single parameter estimation under the locally unbiased condition of Eq. (3), which is weaker than the unbiased condition. In this case, the QCRB is asymptotically attained through adaptive measurements [56, 57]. This attainability does not generally extend to multiple parameter estimations. Specifically, measuring one parameter may affect the precision in the measurement of another parameter, and a trade-off between the MSEM associated with each parameter exists for any collective measurement [8]. Intuitively, any incompatibility among the parameters θ prohibits the simultaneous optimal estimation of all parameters. The RLD QCRB was introduced as the minimum of the weighted sum of the MSES for each parameter under the unbiased condition. Its achievability was shown for Gaussian states [6], qubits [35, 58], and qudit states [50]. However, the RLD QCRB is not always attainable since the optimal estimators derived from the RLD may not correspond to physical POVMs [59].

The problem with saturability of the multiparameter bound was noted by Holevo, who provided the most general quantum extension to the classical Cramér-Rao bound. He introduced a tighter bound, the HCRB $C_H(\theta)$, defined as the maximum among all existing lower bounds for the MSEM of unbiased measurements for the estimation of a set of parameters [60]

$$\text{Tr}[W \Sigma_\theta(\Pi, \check{\theta})] \geq C_H(\theta) \geq \max\{C_S(\theta), C_R(\theta)\}. \quad (7)$$

Helstrom [5] and Holevo [60] demonstrated that $C_H(\theta)$ is attainable if the locally unbiased equality constraints in Eq. (3)

are satisfied. Consider the Hermitian matrix

$$Z(\mathbf{X}) = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] \end{pmatrix}, \quad (8)$$

with $X_j = \int d\theta (\partial_j - \theta_j) \Pi(\theta)$ Hermitian operators. The matrix $Z(\mathbf{X})$ majorises both the inverse of the RLD and SLD Fisher information matrix [35]. Helstrom and Holevo also showed that if $\Sigma_\theta \geq \mathcal{W}$, then $\text{Tr}[\Sigma_\theta] \geq \text{Tr}[\mathcal{W}] + \|\text{Im}\mathcal{W}\|_1$. Together with the RLD CR inequality in Eq. (6), the Holevo bound in Eq. (7) can be written as the solution to the following minimisation [55]

$$\begin{aligned} & \underset{X_1, X_2}{\text{minimize}} && \text{Tr}[\mathcal{W}\text{Re}Z] + \|\mathcal{W}\text{Im}Z\|_1, \\ & \text{subject to} && \text{Tr}[\rho X_j] = 0, \\ & && \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, \end{aligned} \quad (9)$$

where the constraints are the local unbiased conditions of Eq. (3). By minimising over only the first term in the objective function of Eq. (9), we obtain the SLD QCRB [55]

$$C_S(\theta)(\rho, \mathcal{W}) = \text{Tr}[\mathcal{W}[I^S]^{-1}] = \min_{X_1, X_2} \text{Tr}[\mathcal{W}\text{Re}Z]. \quad (10)$$

This shows that the HCRB is indeed a tighter bound than the SLD QCRB, since the second term in Eq. (9) is non-negative [33]. Further, note that the HCRB is not defined explicitly in terms of a closed form for a given statistical model. This is in contrast to the classical case, where the Fisher information can be readily determined from a given statistical model. Recent efforts have focused on determining upper bounds to the HCRB [61–63]. Specifically, the HCRB is upper bounded by a quantity that is twice the SLD-QCRB, such that [61]

$$\max \{C_S(\theta), C_R(\theta)\} \leq C_H(\theta) \leq 2C_S(\theta). \quad (11)$$

In this paper, we provide an analytic solution to the HCRB and provide conditions on when it is tight.

The HCRB is the best asymptotically achievable bound under the conditions stated in Refs [34–36, 50]. Both inequalities in Eq. (7) can be tight [64]. For instance, consider the skew-symmetric matrix $\text{Im}(\text{Tr}[L_j L_k \rho(\theta)])$. When

$$\text{Im}(\text{Tr}[L_j L_k \rho(\theta)]) = 0, \quad (12)$$

we have $C_H(\rho, \theta, \mathcal{W}) = C_S(\rho, \theta, \mathcal{W})$ [8]. This condition is referred to as the weak commutativity criterion [65], and when it is fulfilled the QCRB is a good proxy for the HCRB. In the next section, we show how we can use methods from optimisation theory to address the minimisation over several Hermitian operators in the case where the weak commutativity criterion is not fulfilled.

III. HOLEVO CRAMÉR-RAO BOUND

In this section, we present algorithms to calculate bounds on the Holevo Cramér-Rao bound C_H . We first derive simple

analytic upper and lower bounds for C_H for two parameters in section III A. We show how these bounds are generated by studying the optimisation problem using the method of Lagrange multipliers. This has the advantage of reducing the complexity involved in evaluating bounds on C_H to that of solving two sets of linear equations. In section III B, we focus on deriving lower bounds on C_H for more than two parameters. At the expense of additional analysis, our formalism can be extended to also provide tight analytic solutions to the HCRB. We demonstrate this in section III C, where we provide a complete exposition of analytic bounds on the two-parameter HCRB, and provide conditions for when they bounds are tight.

A. Simple bounds in the two-parameter setting

We first consider the HCRB for two parameters $\theta = (\theta_1, \theta_2)^T$. To obtain simple analytic bounds to the HCRB, we must define the weight matrix \mathcal{W} for the scalar bound. For simplicity, we use the identity weight matrix and determine upper and lower bounds to the two-parameter HCRB using optimisation theory [66]. We want to solve the minimisation in Eq. (9), which is convex but not quadratic. Hence, we first manipulate Eq. (9) into a quadratic form in the variables X_1 and X_2 . Then, such an optimisation problem can be studied analytically using the method of Lagrange multipliers.

Choosing $Y = X_1 + iX_2$, Eq. (9) can be written as an optimisation program (see appendix B)

$$\begin{aligned} & \underset{Y, t}{\text{minimize}} && t, \\ & \text{subject to} && \text{Tr}[Y \rho Y^\dagger] \leq t, \text{Tr}[Y^\dagger \rho Y] \leq t, \\ & && \text{Tr}[\rho Y] = 0, \text{Tr}[\partial_1 \rho Y] = 1, \text{Tr}[\partial_2 \rho Y] = i. \end{aligned} \quad (13)$$

Note that by considering both the real and imaginary parts of the above equality constraints, the actual number of real-valued equality constraints is six, which is consistent with the number of equality constraints corresponding to the minimisation in Eq. (9). Here Y is optimised over all complex matrices of dimension D , and is in general not a Hermitian matrix. By mapping Y and t into a real vector \mathbf{x} , we cast this optimisation program into the standard form of

$$\min_{\mathbf{x}} \{f(\mathbf{x}) : c_i(\mathbf{x}) \leq 0, h_i(\mathbf{x}) = 0\}, \quad (14)$$

where $f(\mathbf{x})$ is a real linear objective function, while $h_i(\mathbf{x})$ and $c_i(\mathbf{x})$ are the corresponding equality and inequality constraint functions that must also be real. Eq. (13) is a convex program, since its equality constraints are linear and its inequality constraints are quadratic and convex. To check whether we can use optimality conditions from optimisation theory, we check whether Slater's constraint qualification holds. This amounts to checking that all the inequality constraints in Eq. (13) can strictly hold. Since t can be arbitrarily large, this indeed is the case. The optimality conditions for a continuous optimisation program are best stated in terms of the Lagrangian of Eq. (13),

given by

$$L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{z}) = f(\mathbf{x}) + \sum_{i=1}^2 \lambda_i c_i(\mathbf{x}) + \sum_{i=1}^6 z_i h_i(\mathbf{x}), \quad (15)$$

where the coefficients $\lambda_i \geq 0$ and $z_i \in \mathbb{R}$ are Lagrange multipliers for the inequality and equality constraints respectively. Since Eq. (14) is a convex program and Slater's constraint qualification holds, the first order Karush-Kuhn-Tucker (KKT) conditions of stationarity, primal and dual feasibility, and complementary slackness are necessary and sufficient [66] to determine the optimality of Eq. (13).

For our problem we have dual variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) = (u, v)^\top$ and $\mathbf{z} = (z_1, \dots, z_6)^\top$, which are vectors of Lagrange multipliers. The primal variables are Y and t , and the Lagrangian is given by

$$L(Y, t, u, v, \mathbf{z}) = t(1 - u - v) - \mathbf{b}^\top \mathbf{z} + u \text{Tr}[Y \rho Y^\dagger] + v \text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger]. \quad (16)$$

Here $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$ is a column vector that encodes the equality constraints in Eq. (13), constructed in Appendix B 2. The operator A is a linear superposition of ρ and its derivatives,

$$A = z_1 A_1 + \dots + z_6 A_6, \quad (17)$$

where

$$\begin{aligned} A_1 &= \frac{1}{2} \rho, & A_4 &= -i A_1, \\ A_2 &= \frac{1}{2} \partial_1 \rho, & A_5 &= -i A_2, \\ A_3 &= \frac{1}{2} \partial_2 \rho, & A_6 &= -i A_3. \end{aligned} \quad (18)$$

Due to the duality principle in optimisation theory [66], we may equivalently view the optimisation by considering the Lagrange dual function $g(\boldsymbol{\lambda}, \mathbf{z}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{z})$ of Eq. (15). Since the Lagrangian L is quadratic in \mathbf{x} , the Lagrange dual can be found analytically by an unconstrained minimisation of the Lagrangian with respect to \mathbf{x} for fixed values of the dual variables $\boldsymbol{\lambda}$ and \mathbf{z} [66]. Due to the structure of the Lagrangian in Eq. (16), the Lagrange dual is never unbounded from below whenever $u + v = 1$. Hence, maximising the Lagrange dual function corresponds to an unconstrained maximisation problem. Since the Lagrange dual is also a quadratic function in terms of its dual variables \mathbf{z} , it can be easily maximised exactly with respect to \mathbf{z} .

Note that our Lagrange dual is not a quadratic function with respect to $\boldsymbol{\lambda} = (u, v)$. To bound C_H , it suffices to evaluate the Lagrangian for feasible values of (u, v) that satisfy $u + v = 1$. Two such values are the boundary values $(u, v) = \{(0, 1), (1, 0)\}$, for which the Lagrangian in Eq. (16) is greatly simplified. For each case, we first determine the stationary point of the resulting Lagrangian with respect to Y , where Y has an implicit dependence on \mathbf{z} , and then perform a maximisation over \mathbf{z} . By evaluating the primal and dual objective functions, we obtain simple analytic two-sided bounds

for C_H . Specifically, an analytic lower bound \mathcal{L} to the HCRB is determined through finding $\mathbf{z} \in \mathbb{R}^6$ that solves

$$2\text{Re}(Q_j)\mathbf{z} + \mathbf{b} = 0, \quad j = 1, 2 \quad (19)$$

where Q_j has the matrix elements

$$[Q_1]_{ik} = \text{Tr}[A_i^\dagger \rho^{-1} A_k] \quad \text{and} \quad [Q_2]_{ik} = \text{Tr}[A_i \rho^{-1} A_k^\dagger]. \quad (20)$$

Details for this are delegated to appendix B. The matrices $\text{Re}(Q_j)$ are full rank when the derivatives $\partial_1 \rho$ and $\partial_2 \rho$ are linearly independent. Armed with these dual variables \mathbf{z} , we collect the result of this optimisation in the following theorem:

Theorem 1. *Let $\partial_1 \rho$ and $\partial_2 \rho$ be linearly independent. With Q_j defined in Eq. (20) and the matrices A_1, \dots, A_6 given in Eq. (18), the HCRB C_H for two parameters satisfies the inequality*

$$\max_{j=1,2} \{l_j\} = \mathcal{L} \leq C_H \leq \mathcal{U} = \min_{j=1,2} \{\max\{l_j, m_j\}\},$$

where

$$l_j = \frac{1}{4} \mathbf{b}^\top \text{Re}(Q_j)^{-1} \mathbf{b}, \quad (21)$$

$$m_j = \sum_{a,b=1}^6 \text{Tr}[\rho^{-2} A_a^\dagger \rho A_b] z_{a,j} z_{b,j}. \quad (22)$$

and

$$z_{a,j} = -\frac{1}{2} \left([\text{Re}(Q_j)^{-1}]_{a2} + [\text{Re}(Q_j)^{-1}]_{a6} \right). \quad (23)$$

For a detailed proof of this theorem, consult appendices B 2-B 5. Theorem 1 gives a simple procedure for finding analytic upper and lower bounds to two-parameter HCRB. Notice that the complexity of determining these bounds are commensurate with linear equation solvers that are used in determining the Lagrange dual variables. This makes these bounds readily accessible for general two-parameter applications. Fig. 1 shows the pseudocode for this procedure.

Next, we establish how to construct the observables that saturate these bounds. Specifically, there are two choices for Y that minimise the Lagrangian in Eq. (16), corresponding to the two choices for (u, v) :

$$(u, v) = (1, 0) : \quad Y = -A^\dagger \rho^{-1}, \quad (24)$$

$$(u, v) = (0, 1) : \quad Y = -\rho^{-1} A^\dagger, \quad (25)$$

that correspond to the choice Q_1 and Q_2 , respectively. Then, using Eq. (17) and the optimised values for \mathbf{z} , we construct the analytic form for the observables X_1 and X_2 . Since the matrices $Q_j, j \in \{1, 2\}$ are only 6 dimensional matrices, determining $\text{Re}(Q_j)$ is easy and hence it is straightforward to find analytic bounds to the two-parameter HCRB. The procedure is shown algorithmically in Fig. 1.

Finally, we note that if our lower bound to C_H is strictly larger than C_S , then we know that the skew-symmetric matrix $\text{Im}(\text{Tr}[L_j L_k \rho(\boldsymbol{\theta})])$ cannot be equal to zero, and the weak commutativity criterion does not hold.

Algorithm 1 $\{\mathcal{L}, \mathcal{U}, X_1, X_2\} = \text{SIMPLE}(\rho, \partial_1\rho, \partial_2\rho)$

```

1:  $A_1 \leftarrow \rho/2$ 
2:  $A_2 \leftarrow \partial_1\rho/2$ 
3:  $A_3 \leftarrow \partial_2\rho/2$ 
4:  $A_4 \leftarrow -iA_1$ 
5:  $A_5 \leftarrow -iA_2$ 
6:  $A_6 \leftarrow -iA_3$ 
7:  $\mathbf{b} \leftarrow [0, 1, 0, 0, 0, 1]^T$ 
8:  $Q_1, Q_2 \leftarrow$  size 6 zeros matrix
9: for all  $j, k = 1$  to 6 do
10:    $[Q_1]_{jk} = \text{Tr}[A_j^\dagger \rho^{-1} A_k]$ 
11:    $[Q_2]_{jk} = \text{Tr}[A_j \rho^{-1} A_k^\dagger]$ 
12: Solve for  $\mathbf{y} = [y_1, \dots, y_6]^T$  in  $2\text{Re}(Q_1)\mathbf{y} + \mathbf{b} = 0$ 
13: Solve for  $\mathbf{z} = [z_1, \dots, z_6]^T$  in  $2\text{Re}(Q_2)\mathbf{z} + \mathbf{b} = 0$ 
14:  $v_1 \leftarrow -\mathbf{y}^T \text{Re}(Q_1)\mathbf{y} - \mathbf{b}^T \mathbf{y}$ 
15:  $v_2 \leftarrow -\mathbf{z}^T \text{Re}(Q_2)\mathbf{z} - \mathbf{b}^T \mathbf{z}$ 
16: if  $v_1 > v_2$  then
17:    $\mathcal{L} \leftarrow v_1$ 
18:    $A \leftarrow y_1 A_1 + \dots + y_6 A_6$ 
19:    $Y \leftarrow -A^\dagger \rho^{-1}$ 
20:    $\mathcal{U} \leftarrow \max\{\text{Tr}[Y \rho Y^\dagger], \text{Tr}[Y^\dagger \rho Y]\}$ 
21: else
22:    $\mathcal{L} \leftarrow v_2$ 
23:    $A \leftarrow z_1 A_1 + \dots + z_6 A_6$ 
24:    $Y \leftarrow -\rho^{-1} A^\dagger$ 
25:    $\mathcal{U} \leftarrow \max\{\text{Tr}[Y \rho Y^\dagger], \text{Tr}[Y^\dagger \rho Y]\}$ 
26:  $X_1 \leftarrow (Y + Y^\dagger)/2$ 
27:  $X_2 \leftarrow (Y - Y^\dagger)/(2i)$ 
28: Return  $\{\mathcal{L}, \mathcal{U}, X_1, X_2\}$ 

```

FIG. 1: Pseudocode to determine simple bounds to the two-parameter HCRB and its associated optimal measurement observables X_1 and X_2 . Note that this algorithm depends only on the state ρ and its two derivatives $\partial_1\rho$ and $\partial_2\rho$.

B. Lower bound in the multi-parameter setting

For more than two parameters, we can also use the method of Lagrange multipliers to bound the HCRB. However, this method is considerably more involved than the two-parameter case. In the two-parameter case, we could obtain a simple quadratic expression for $\text{ReTr}[Z] + \|\text{ImZ}\|_1$ that appears in the objective function of Eq. (9). However for the corresponding generalisation to multiple parameters, $\text{ReTr}[Z] + \|\text{ImZ}\|_1$ is no longer a quadratic form in the variables X_j . For example, for three parameters Z takes the form

$$Z(X) = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] & \text{Tr}[\rho X_1 X_3] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] & \text{Tr}[\rho X_2 X_3] \\ \text{Tr}[\rho X_3 X_1] & \text{Tr}[\rho X_3 X_2] & \text{Tr}[\rho X_3^2] \end{pmatrix}. \quad (26)$$

The trace norm of ImZ is related to the eigenvalues of ImZ , and the eigenvalues of a 3×3 matrix involve a cubic equation. This renders evaluating the trace norm incompatible with our methodology. To address this, we obtain a lower bound

to $\|\text{ImZ}\|_1$ that allows $\text{ReTr}[Z] + \|\text{ImZ}\|_1$ to be written as a quadratic form. This yields an optimisation problem whose optimal value is a lower bound to the HCRB, and which is given by

$$\min\{t : \text{Tr}[\rho X_j] = 0, \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, V_\alpha \leq t\}. \quad (27)$$

The minimisation is performed over Hermitian matrices X_1, \dots, X_d and t . The indices j, k run over $\{1, \dots, d\}$. The inequality constraint V_α is a function of a binary string α such that

$$V_\alpha = \frac{1}{2} \sum_{j=1}^d \text{Tr}[(X_j + (-1)^{\alpha_j} i X_{j+1}) \rho (X_j + (-1)^{\alpha_j} i X_{j+1})^\dagger], \quad (28)$$

with $X_{d+1} = X_1$. The inequality constraints V_α arise from the structure of our lower bound on the trace norm of ImZ (see appendix C). By substituting $Y_j = X_j + i X_{j+1}$, we can write the matrices X_j in terms of the matrices Y_j , as before. We next interpret the Y_j as arbitrary complex matrices of size n , and impose Hermiticity conditions for the corresponding X_j matrices.

The Lagrangian of such an optimisation problem is a function of the complex matrices $\{Y_1, \dots, Y_d\}$, and also a function of its Lagrange multipliers. Its Lagrange multipliers are given by the non-negative multipliers $\mathbf{v} \in \mathbb{R}^{2^d}$ for the inequality constraints, $\mathbf{z} \in \mathbb{R}^{d(d+1)}$ for the equality constraints, and Hermitian multipliers ξ_1, \dots, ξ_d for the Hermitian constraints. Most importantly, the inequality constraints can be satisfied strictly, so Slater's constraint qualification holds, and we can use the KKT to determine the optimality conditions for Eq. (28). We minimise the Lagrangian constructed from the optimisation problem in Eq. (27). Since the Lagrangian is a convex quadratic form in the variables Y_1, \dots, Y_d , it can be minimised exactly. When this is done, we obtain the Lagrange dual function, which only depends on the Lagrange multipliers \mathbf{v}, \mathbf{z} and ξ_1, \dots, ξ_d . The Lagrange dual function always gives a lower bound for the primal optimisation problem.

While the Lagrange dual is quadratic in \mathbf{z} and ξ_1, \dots, ξ_d , it is not quadratic in \mathbf{v} . By minimising the Lagrangian over t and using the KKT conditions, we conclude as before that the sum of the components in \mathbf{v} is 1. We obtain a lower bound for the Lagrange dual by maximising over a discrete set of feasible Lagrange multipliers \mathbf{v} , which corresponds to the tightness of the constraints $V_\alpha \leq t$. Thus, we created a quadratic optimisation problem for three or more parameters that leads to a lower bound on C_H . However, there is no guarantee that this lower bound is tight.

Next, we study the Lagrange dual function. By carefully choosing \mathbf{v} , the Lagrangian is quadratic in $\{Y_1, \dots, Y_d\}$ and can be minimised individually for each Y_j . The coefficients for Y_j in the Lagrangian are given by Γ_j , where

$$\Gamma_j = \sum_{k=1}^d T_{k,j} \left(\sum_{l=0}^d Z_{l,k} \rho_l + i \xi_k \right), \quad (29)$$

where $\rho_0 = \rho$, $\rho_j = \partial_j \rho$ for $j = 1, 2$, and $T_{k,j}$ are matrix elements that relate the Y_j to the X_k . Then the optimal value

Algorithm 2 $v = g_\alpha(Z, \xi_1, \dots, \xi_d)$

```

1:  $S \leftarrow$  size  $d$  zeros matrix
2: for all  $j = 1$  to  $d$  do
3:    $S_{jj} = 1$ 
4:   if  $j \neq d$  then
5:      $k \leftarrow j + 1$ 
6:   else
7:      $k \leftarrow 1$ 
8:    $S_{jk} = i$ 
9:  $T \leftarrow S^{-1}$ 
10:  $v \leftarrow -\sum_{j=1}^d [Z]_{jj}$ 
11: for all  $j = 1$  to  $d$  do
12:    $\Gamma_j \leftarrow \sum_{k=1}^d T_{k,j} \left( i\xi_k + \sum_{l=0}^d [Z]_{l,k} \rho_l \right)$ 
13:   if  $\alpha_j = 0$  then
14:      $v \leftarrow v - \text{Tr}[\Gamma_j \rho^{-1} \Gamma_j^\dagger] / 2$ 
15:   else
16:      $v \leftarrow v - \text{Tr}[\Gamma_j^\dagger \rho^{-1} \Gamma_j] / 2$ 
17: Return  $v$ 

```

FIG. 2: Pseudocode to generate the Lagrange dual functions defined in Eq. (30).

for the Lagrange multipliers can be obtained by maximising the Lagrange dual functions

$$g_\alpha = -\sum_{j=1}^d z_{j,j} - \sum_{j=1}^d \frac{\delta_{0,\alpha_j} \text{Tr}[\Gamma_j \rho^{-1} \Gamma_j^\dagger] + \delta_{1,\alpha_j} \text{Tr}[\Gamma_j^\dagger \rho^{-1} \Gamma_j]}{2}. \quad (30)$$

with respect to the scalar variables $z_{j,k}$ and the Hermitian variables ξ_j . Our lower bound to C_H in the multi-parameter setting is then given by the following theorem.

Theorem 2. *Let $d \geq 3$, $\mathbf{z} \in \mathbb{R}^{d(d+1)}$ and ξ_1, \dots, ξ_d are Hermitian matrices. Then*

$$C_H \geq \max_{\alpha \in \{0,1\}^d} \max_{\mathbf{z}, \xi_1, \dots, \xi_d} g_\alpha,$$

where g_α is given by Eq. (30).

This optimisation problem can be solved exactly using a single step of Newton's method. It requires the input state ρ and its derivatives $\partial_j \rho$. The algorithm to implement this lower bound is illustrated in Fig. 3.

C. Tight two-parameter bounds

Notice that for Theorem 1, we constrained the values of the Lagrange multiplier u to two values. This does not provide the most general case and as a result, the analysis can generate observables that are not always optimal. That is, the corresponding upper and lower bounds are not always tight. By lifting this restriction, we expand the analysis to explore the full generality of our formalism to generate tight bounds to the estimation of incompatible observables. As we observe in this section, this is necessary to develop an intuition into

Algorithm 3 $\{\mathcal{L}, X_1, \dots, X_d\} = \text{MULTI}(\rho, \partial_1 \rho, \dots, \partial_d \rho)$

```

1:  $\rho_0 \leftarrow \rho$ 
2:  $S \leftarrow$  size  $d$  zeros matrix
3: for all  $j = 1$  to  $d$  do
4:    $\rho_j \leftarrow \partial_j \rho$ 
5:    $S_{jj} = 1$ 
6:   if  $j \neq d$  then
7:      $k \leftarrow j + 1$ 
8:   else
9:      $k \leftarrow 1$ 
10:   $S_{jk} = i$ 
11:  $T \leftarrow S^{-1}$ 
12: for all binary vectors  $\alpha = (\alpha_1, \dots, \alpha_d)$  do
13:   Solve  $\max g_\alpha(Z, \xi_1, \dots, \xi_d)$  w.r.t. the real matrix  $Z$ .
       $Z$ : rows labelled from 0 to  $d$ , columns labelled from 1 to  $d$ 
       $\xi_1, \dots, \xi_d$ : complex size  $d$  Hermitian matrices
14:   Let  $(Z_{\text{opt}}, \xi_{\text{opt},1}, \dots, \xi_{\text{opt},d})$  denote the optimal solution
15:    $v_\alpha \leftarrow g_\alpha(Z_{\text{opt}}, \xi_{\text{opt},1}, \dots, \xi_{\text{opt},d})$  (See Algorithm 2)
16:  $\mathcal{L} \leftarrow 0$ 
17: for all binary vectors  $\alpha = (\alpha_1, \dots, \alpha_d)$  do
18:   if  $v_\alpha > \mathcal{L}$  then
19:      $\mathcal{L} \leftarrow v_\alpha$ 
20:     for all  $j = 1$  to  $d$  do
21:        $\Gamma_j \leftarrow \sum_{k=1}^d T_{k,j} \left( i\xi_{\text{opt},k} + \sum_{l=0}^d [Z_{\text{opt}}]_{l,k} \rho_l \right)$ 
22:       if  $\alpha_j = 0$  then
23:          $Y_j \leftarrow -\Gamma_j^\dagger \rho^{-1}$ 
24:       else
25:          $Y_j \leftarrow -\rho^{-1} \Gamma_j^\dagger$ 
26:     for all  $k = 1$  to  $d$  do
27:        $X_k \leftarrow \sum_{j=1}^d T_{k,j} Y_j$ 
28: Return  $\{\mathcal{L}, X_1, \dots, X_d\}$ 

```

FIG. 3: Pseudocode to generate a lower bound to the HCRB for more than two parameters.

multiparameter quantum estimation that is captured by the construction of the HCRB. To achieve this, we revisit the two-parameter scenario. Specifically, for fixed u , we minimise the Lagrangian and find the optimal observables that attain these stationary points. In doing so, for every feasible value of u , we obtain an upper and lower bound on the HCRB. Since the lower bound is a concave and smooth function, then optimisation theory guarantees a solution to both the HCRB and the observables that attain it.

Recall the Lagrangian in Eq. (16), which for $u + v = 1$ becomes

$$\begin{aligned} \mathcal{L}(Y, u, \mathbf{z}) = & -\mathbf{b}^\top \mathbf{z} + u \text{Tr}[Y \rho Y^\dagger] + (1-u) \text{Tr}[Y^\dagger \rho Y] \\ & + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger]. \end{aligned} \quad (31)$$

As we show in appendix D, this Lagrangian is minimised when Y is chosen such that

$$uY\rho + (1-u)\rho Y = -A^\dagger. \quad (32)$$

Notice that Y , and hence the optimal observables, is the solution to a Sylvester equation. When $\text{Im}Z = 0$, we know that the observables X_1 and X_2 commute and the weak commutativity criterion is preserved. Then, Eq. (32) reduces to solving a

Algorithm 4 $\{\mathcal{L}, \mathcal{U}, X_1, X_2\} = \text{MASTER}(\rho, \partial_1 \rho, \partial_2 \rho)$

```

1: Find  $u \in [0, 1]$  that maximises
    $\mathcal{L}_u$  that is returned from  $\text{HCRBU}(u, \rho, \partial_1 \rho, \partial_2 \rho)$ 
2:  $\{\mathcal{L}_u, \mathcal{U}_u, Y\} = \text{HCRBU}(u, \rho, \partial_1 \rho, \partial_2 \rho)$ 
3:  $X_1 \leftarrow (Y + Y^\dagger)/2$ 
4:  $X_2 \leftarrow (Y - Y^\dagger)/(2i)$ 
5: Return  $\{\mathcal{L}_u, \mathcal{U}_u, X_1, X_2\}$ 

Procedure:  $\{\mathcal{L}_u, \mathcal{U}_u, Y\} = \text{HCRBU}(u, \rho, \partial_1 \rho, \partial_2 \rho)$ 
1:  $D \leftarrow \text{size of } \rho$ 
2:  $\{p_j, |e_j\rangle : j = 1, \dots, D\} \leftarrow \text{eigenvalues and eigenvectors of } \rho$ 
3:  $|e_j\rangle \leftarrow |e_j\rangle \langle e_j|$ 
4:  $\gamma_1 \leftarrow -\mathbf{1}_D/2$ 
5:  $\gamma_2 \leftarrow \sum_{j,k=1}^D (up_k + (1-u)p_j)^{-1} \langle e_j | (\partial_1 \rho/2) | e_k \rangle |e_j\rangle \langle e_k|$ 
6:  $\gamma_3 \leftarrow \sum_{j,k=1}^D (up_k + (1-u)p_j)^{-1} \langle e_j | (\partial_2 \rho/2) | e_k \rangle |e_j\rangle \langle e_k|$ 
7:  $G_1 \leftarrow \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2 \rho \gamma_2^\dagger] & \text{Tr}[\gamma_2 \rho \gamma_3^\dagger] \\ 0 & \text{Tr}[\gamma_3 \rho \gamma_2^\dagger] & \text{Tr}[\gamma_3 \rho \gamma_3^\dagger] \end{pmatrix}$ 
8:  $G_2 \leftarrow \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2^\dagger \rho \gamma_2] & \text{Tr}[\gamma_2^\dagger \rho \gamma_3] \\ 0 & \text{Tr}[\gamma_3^\dagger \rho \gamma_2] & \text{Tr}[\gamma_3^\dagger \rho \gamma_3] \end{pmatrix}$ 
9:  $G_3 \leftarrow \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\partial_1 \rho \gamma_2] & \text{Tr}[\partial_1 \rho \gamma_3] \\ 0 & \text{Tr}[\partial_2 \rho \gamma_2] & \text{Tr}[\partial_2 \rho \gamma_3] \end{pmatrix}$ 
10:  $Q_k \leftarrow \begin{pmatrix} G_k & (-1)^k i G_k \\ (-1)^{k-1} i G_1 & G_k \end{pmatrix}, k = 1, 2.$ 
11:  $Q_3 \leftarrow \frac{1}{2} \begin{pmatrix} G_3 + G_3^* & i(G_3 - G_3^*) \\ -i(G_3 - G_3^*) & G_3 + G_3^* \end{pmatrix}$ 
12:  $Q \leftarrow u \text{Re}(Q_1) + (1-u) \text{Re}(Q_2) + \text{Re}(Q_3)$ 
13: trigger  $\leftarrow 0$ 
14: KernelDim  $\leftarrow 6 - \text{rank}(Q)$ 
15:  $\{\mathbf{ker}_{\text{idx}}\}_{\text{idx}=1:\text{KernelDim}} \leftarrow \text{orthonormal basis of } Q\text{'s kernel.}$ 
16: for all idx = 1 : KernelDim do
17:   trigger  $\leftarrow \text{trigger} + \|\mathbf{b}^\top \mathbf{ker}_{\text{idx}}\|_1$ 
18: if trigger > 0 then
19:    $\mathcal{L}_u, \mathcal{U}_u \leftarrow \infty$ 
20:    $Y \leftarrow \text{matrix of NaN of size } D$ 
21: else
22:    $\mathbf{b} \leftarrow [0, 1, 0, 0, 0, 1]^\top$ 
23:   if rank( $Q$ ) = 6 then
24:     Solve for  $\mathbf{z} = [z_1, \dots, z_6]^\top$  in  $2Q\mathbf{z} = \mathbf{b}$ 
25:   else
26:     Find  $\mathbf{z} = [z_1, \dots, z_6]^\top$  that maximises  $-\mathbf{b}^\top \mathbf{z} + \mathbf{z}^\top Q \mathbf{z}$ 
27:    $\mathcal{L}_u \leftarrow -\mathbf{b}^\top \mathbf{z} + \mathbf{z}^\top Q \mathbf{z}$ 
28:    $\mathcal{U}_u \leftarrow \max(\mathbf{z}^\top Q_1 \mathbf{z}, \mathbf{z}^\top Q_2 \mathbf{z})$ 
29:    $Y \leftarrow z_1 \gamma_1 + z_2 \gamma_2 + z_3 \gamma_3 + i(z_4 \gamma_1 + z_5 \gamma_2 + z_6 \gamma_3)$ 
30: Return  $\{\mathcal{L}_u, \mathcal{U}_u, Y\}$ 

```

FIG. 4: Master algorithm to generate the analytic form of the two-parameter HCRB.

Lyapunov equation, where Y generates the well explored SLD. This corresponds to

$$|\text{Tr}[\rho X_1 X_2] - \text{Tr}[\rho X_2 X_1]| = 0, \quad (33)$$

for the optimal X_1 and X_2 , which recovers the weak commutativity condition. In this way, the solution to Y in Eq. (32) provides the most general definition to quantum logarithmic derivatives for multiple incompatible parameters θ . From our

definition of A , it defines exactly how the optimal observables depend on dynamics in both parameters.

Similar to analytic solutions to the SLD, given the spectral decomposition of the state $\rho = \sum_j p_j |e_j\rangle \langle e_j|$, we can analytically solve Eq. (32) to obtain the Y that minimises the Lagrangian:

$$Y = - \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | A^\dagger | e_k \rangle |e_j\rangle \langle e_k|. \quad (34)$$

By taking the trace, it is clear to see that Y has a zero expectation value. This recovers the unbiasedness condition on the observables as required. Using $Y = X_1 + iX_2$ and comparing with Eq. (34), we arrive at an analytic solution for the observables that saturate the HCRB

$$X_1 = -\frac{1}{2} \sum_{j,k} \frac{\langle e_j | z_1 \rho + z_2 \rho_1 + z_3 \rho_2 | e_k \rangle}{(up_k + (1-u)p_j)} |e_j\rangle \langle e_k|, \quad (35)$$

$$X_2 = -\frac{1}{2} \sum_{j,k} \frac{\langle e_j | z_1 \rho + z_4 \rho_1 + z_5 \rho_2 | e_k \rangle}{(up_k + (1-u)p_j)} |e_j\rangle \langle e_k|, \quad (36)$$

where z 's are the optimal choice for the Lagrange multipliers. This shows exactly how each observable depend on the dynamics of each parameter.

With access to the spectral decomposition of ρ , we can also find analytic expressions to the HCRB. The master algorithm in Fig. 4 concisely clarifies this procedure. With this procedure, theorem 3 concretely demonstrates how to construct tight bounds on the HCRB, and is central to our result.

Theorem 3. Let Q_1 , Q_2 and Q_3 be matrices defined in the Master algorithm in Fig. 4, and let $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$ be a column vector. Let $0 < u < 1$ and $Q = u \text{Re} Q_1 + (1-u) \text{Re} Q_2 + \text{Re} Q_3$ be a negative definite matrix. Then when Q is full rank for all $0 < u < 1$, the HCRB is bounded through $\mathcal{L}_u \leq C_H \leq \mathcal{U}_u$ where

$$\mathcal{L}_u = -\frac{1}{4} \mathbf{b}^\top Q^{-1} \mathbf{b} \quad (37)$$

$$\mathcal{U}_u = \frac{1}{4} \max \{ \mathbf{b}^\top Q^{-1} Q_1 Q^{-1} \mathbf{b}, \mathbf{b}^\top Q^{-1} Q_2 Q^{-1} \mathbf{b} \} \quad (38)$$

with equality on both sides attained at the stationary point of the lower bound with respect to u when $\frac{d}{du}(\mathbf{b}^\top Q^{-1} \mathbf{b}) = (\mathbf{b}^\top Q^{-1} \frac{dQ}{du} Q^{-1} \mathbf{b}) = 0$.

We refer the reader to Appendix D for a complete proof of this theorem. For applications where the spectral decomposition of the state is not known, the Sylvester equation (34) can be efficiently solved numerically using a variant of the Bartel-Stewart algorithm [67].

Before concluding, we clarify an important subtlety. Our analysis assumes that the probe state is fixed. However, there are multi-parameter sensing applications that permit full control over the probe states used. In this case, it is possible to extend our formalism to determine the optimal probe state. To see how, note that the HCRB is a bi-convex function of the probe state ρ and the observable X . We have already determined the optimal observable corresponding to a chosen state:

$C_{H,\rho}(X) = C_H(\rho, X)$. Conversely, fixing X amounts to solving a convex problem in ρ to determine the optimal state corresponding to the choice in X : $C_{H,X}(\rho) = C_H(\rho, X)$. Based on this, we can implement an efficient iterative bi-convex program that alternatively updates the state and optimal observables by fixing one and solving the corresponding convex optimisation problem [68].

IV. APPLICATIONS

Quantum metrology has applications in both spin and bosonic systems. We demonstrate the broad applicability of our results by showing how our bounds work in each of the two settings. First, for spin systems, one natural problem to consider is that of estimating the different components of a magnetic field. When the total magnetic field is known, there are only two independent components of a magnetic field to estimate, and such a problem can be tackled directly using our analytical approach for two-parameter estimation. In particular, our simple approach using the Algorithm in Fig. 1 already gives interesting insights into the problem of quantum magnetometry on various types of noisy probe states.

Secondly, for bosonic systems a key obstacle in determining the ultimate precision limits on parameter estimation is the infinite dimension of such systems. We show that using our analytical approach, this obstacle can be overcome. Specifically, we use our tight analytical bounds (algorithm 4) to calculate the precision bounds on estimating the incompatible components of a logical Bloch vector of a pure bosonic codestate when mixed with a thermal state.

A. Magnetometry

We use our simple two-parameter bounds to consider magnetic field sensing, which has important technological applications in navigation, position tracking, and imaging [69]. We apply our method of finding the HCRB to the estimation of a magnetic field $\mathbf{B} = (B_x, B_y, B_z)$ in three dimensions. Quantum magnetometry is an important application of quantum metrology, and is essential for detecting defects and realising compact magnetic resonance imaging scanners [70]. Estimating each component individually allows us to attain the quantum limit [23], and this has been demonstrated in several studies [71, 72]. However, in many practical applications, knowledge of multiple parameters is required simultaneously, and we must consider joint estimation strategies.

The three parameters of interest $\theta = (\theta_1, \theta_2, \theta_3)^\top$ appear in the single spin Hamiltonian $\hat{H}_j(\theta) = \theta \cdot \mathbf{S}_j$, where \mathbf{S}_j is the spin operator for the j^{th} spin. Local depolarising noise, described by the single spin CPTP map

$$\mathcal{D}_g[\rho] = (1 - g)\rho + g\frac{\mathbb{1}_2}{2}, \quad (39)$$

provides a general description for a noisy environment, where g denotes the depolarisation magnitude and takes values between 0 and 1. The parameters θ are imprinted on the probe

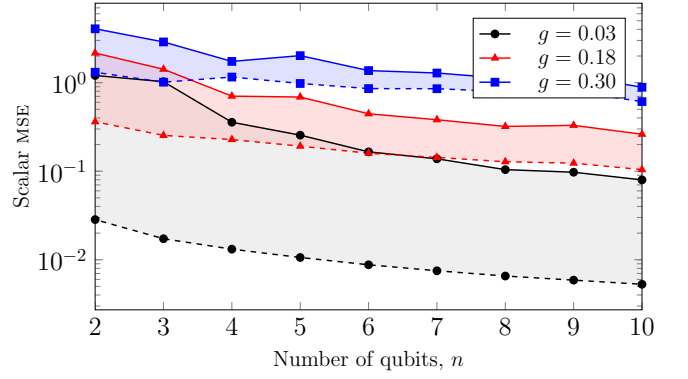


FIG. 5: Estimation precision of two directional components of the a magnetic field B_x and B_y with increasing number of spins in a depolarising environment parameterised by g , using modified 3D-GHZ states. The corresponding Hamiltonian is $\sum_{j=1}^n \hat{H}_j(\theta)$ and has no interactions terms. The solid lines and dashed lines represent upper and lower bounds for the HCRB as given by theorem 1.

state via the unitary evolution $\hat{U} = \exp[-i\hat{H}_j(\theta)]$. For our example, we assume that the magnetic field in the z -direction is known, and we therefore wish to estimate the two parameters B_x and B_y . We choose an identity weight matrix to equally prioritise each parameter into a weighted scalar mean square error. We consider three families of n -spin probe states, namely the traditional GHZ states for single-parameter estimation, the modified 3D-GHZ states introduced by Baumgratz and Datta [13], and the GNU states introduced by Ouyang in the context of quantum error correction [73].

First, the 3D-GHZ state can be written as

$$|\psi_n^{3\text{D-GHZ}}\rangle = \frac{1}{\mathcal{N}} \sum_{j=1}^3 (|\phi_j^+\rangle^{\otimes n} + |\phi_j^-\rangle^{\otimes n}), \quad (40)$$

where n is the total number of spins, \mathcal{N} is the normalisation constant of the state and $|\phi_j^\pm\rangle$ are the eigenvectors corresponding to the ± 1 eigenvectors of the j^{th} spin matrix. The evolved state then becomes

$$\rho(\theta) = \hat{U}(\theta) \mathcal{D}_g^{\otimes n} [|\psi_n^{3\text{D-GHZ}}\rangle \langle \psi_n^{3\text{D-GHZ}}|] \hat{U}(\theta)^\dagger. \quad (41)$$

We illustrate how the HCRB and the QCRB change with the number of probe qubits n for different depolarising channel strengths g in Fig. 5. We observe that the HCRB is indeed tighter than the QCRB. Both variance bounds increase with an increasing depolarising probability of the depolarising channel, as expected. The 3D-GHZ state attains the Heisenberg precision scaling for the noiseless case.

Second, we consider the class of GNU states that are robust to a constant amount of erasure and dephasing [74]:

$$|\varphi_1\rangle = \frac{1}{2} \sum_{j=0}^2 \sqrt{\binom{2}{j}} |D_{Gj}^n\rangle, \quad (42)$$

Here, for every $w = 0, \dots, n$, the Dicke state $|D_w^n\rangle$ is a uniform superposition over all computation basis states $|x_1\rangle \otimes \dots \otimes |x_n\rangle$

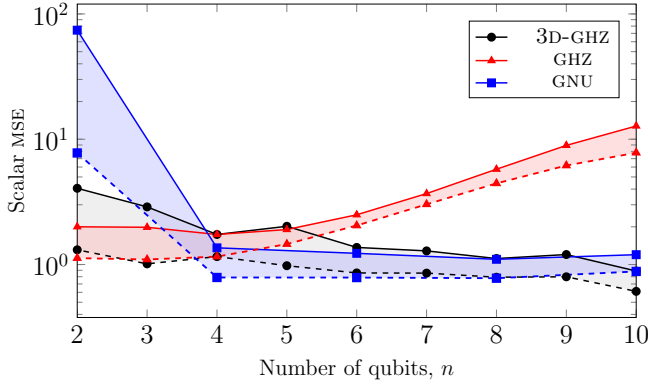


FIG. 6: Using theorem 1, we depict two-sided bounds on the HCRB for estimating the elements of a magnetic field in a depolarising channel. All plots with depolarising strength $g = 0.3$. This plot compares the performance of 3D-GHZ states with n -qubit GHZ states, and the permutation-invariant GNU states. We observe an interesting crossover point between the tightest bound generated by the n -qubit GHZ states and 3D-GHZ states, with increasing number of qubits. The GNU states are defined over the even number of qubit numbers generates the lowest Holevo bound for small number of qubits.

with Hamming weight w . Since $n = 2G$, where G is related to the number of bit-flip errors that can be corrected, we present results for the GNU states for which n is even. These are shown in Fig. 6, and compared with traditional GHZ states and 3D-GHZ states. The traditional GHZ states give a worse estimation for larger qubit number at constant depolarisation rate, as is well-known. The GNU states perform similarly to the 3D-GHZ states.

Finally, we use our formalism to determine the optimal n -qubit observables X_1 and X_2 that attain the HCRB using the 3D-GHZ states. Unlike the single qubit estimation case, analytic solutions to these observables are challenging and the dimension of X_j scales as 2^n . Instead, we numerically determine their structure, shown in Fig. 7. We plot the real and imaginary parts of the matrices X_1 and X_2 , and the Hermiticity of the observables is clearly observed.

B. Bosonic quantum codes

Our formalism for the HCRB allows performance characterisation of fault tolerant quantum codes in the context of quantum metrology. In this section, we apply our master algorithm in Fig. 4 to explore how bosonic error-correcting codes can improve characterisation of logically encoded states in the presence of noise.

Susceptible quantum information can be safeguarded from decoherence by storing it in quantum error-correcting codes (QECC), which in the case of continuous variable (bosonic) quantum systems are subspaces of infinite-dimensional Hilbert space. The working principle of QECC is to project states with errors with high probability onto correctible subspaces labelled by the error syndromes, and dynamically evolve the projected state back to the original code space. When these

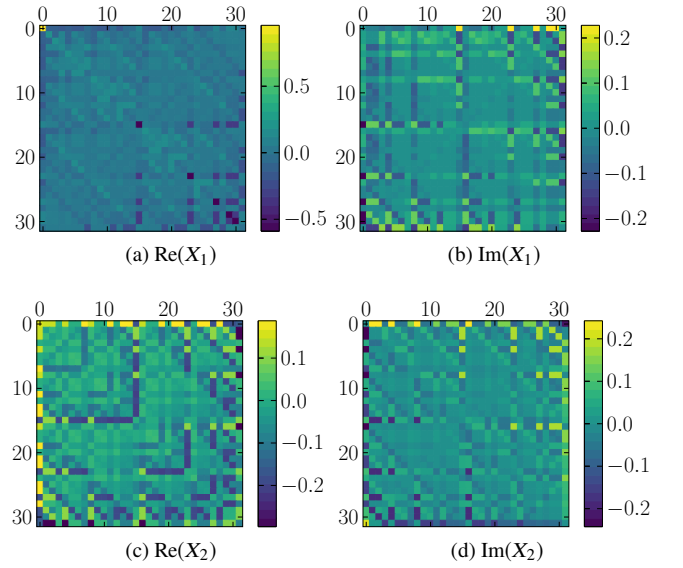


FIG. 7: Using Algorithm 1, we obtain a heatmap for the measurement observables X_1 and X_2 for the depolarised five-qubit 3D-GHZ state, under a depolarising strength $g = 0.3$. We plot the real and imaginary parts separately. The Hermiticity of these observables is clearly illustrated.

codes are well-chosen, they can correct against errors that are introduced in physically realistic noise models. While bosonic codes on multiple-modes that correct against displacement errors [75, 76] and photon loss [77–80] exist, a key attraction of bosonic codes is that they can be used even on a single mode. For example, to protect codes against photon loss and phase errors on a single-mode, one can use codes gapped in the Fock basis [81, 82], or a single-mode GKP code for displacement errors [75]. For a complete exposition of fault tolerant quantum computing and error correcting codes, the reader is directed to references [83–85].

We focus on bosonic codes, which, unlike two-level systems, has infinite energy levels per mode. We define a logically encoded state that is parameterised by the coordinates $\theta = (x, \phi)^T$ with $x = \cos(\theta/2)$. We evaluate the HCRB for this bivariate estimation scheme by using theorem 3 to evaluate upper and lower bounds to the HCRB for fixed u . We then tune the parameters of the binomial codes to effect improvements to estimates of θ . We consider binomial codes that protect codewords against number-shift and phase errors. In particular, we analyse the ultimate limits of estimating the complex coefficients of a pure binomial codestate in the presence of thermal noise.

The logical codewords for the binomial code are supported

on a bounded number of Fock states through

$$|0_L\rangle = \sum_{\substack{j \geq 0 \\ j \text{ even}}} 2^{-\frac{(n-1)}{2}} \sqrt{\binom{n}{j}} |Gj\rangle, \quad (43)$$

$$|1_L\rangle = \sum_{\substack{j \geq 0 \\ j \text{ odd}}} 2^{-\frac{(n-1)}{2}} \sqrt{\binom{n}{j}} |Gj\rangle, \quad (44)$$

where $G, n \in \mathbb{R}$ are related to the number of correctable number-shift and phase errors respectively [92]. From Ref. [81, Eq (7)], one requires $G \geq G_{\text{bin}} + L_{\text{bin}} + 1$ to correct G_{bin} gain and L_{bin} loss errors, and $n - 1 \geq \max\{L_{\text{bin}}, G_{\text{bin}}, 2D_{\text{bin}}\}$ to correct D_{bin} phase errors. For fixed G and n , we construct the logical state $\rho_L = |\psi_L\rangle\langle\psi_L|$ with

$$|\psi_L\rangle = x|0_L\rangle + \sqrt{1-x^2}e^{i\phi}|1_L\rangle, \quad (45)$$

where $x \in [-1, 1]$ and $\phi \in \mathbb{R}$. In the noisy scheme, we thermalise this logical pure state through $\rho = \lambda_{\text{th}}\rho_{\text{th}} + (1 - \lambda_{\text{th}})\rho_L$, where

$$\rho_{\text{th}} = \frac{1}{1 - e^{-\beta}} \sum_{k=0}^{\infty} e^{-k\beta} |k\rangle\langle k|, \quad (46)$$

is a thermal state with temperature β . Since $\partial_j \rho = (1 - \lambda_{\text{th}})\partial_j \rho_L$, $j = \{x, \phi\}$ and ρ_L is only supported on the Fock states $|0\rangle, |G\rangle, \dots, |Gn\rangle$, the state derivatives $\partial_j \rho$ are only supported on the Fock states $|0\rangle, |G\rangle, \dots, |Gn\rangle$. Using this property, we can determine that in the calculation of the HCRB, we need only consider the evaluation of ρ on the support of the Fock states $|0\rangle, |G\rangle, \dots, |Gn\rangle$. Denoting such a state as τ , we can write

$$\tau = \sum_{j,k=0}^n |j\rangle\langle k| \langle Gj|\rho|Gk\rangle, \quad (47)$$

and observe that its has spectral decomposition

$$\tau = \sum_{k=0}^n t_k |\tau_k\rangle\langle\tau_k| \quad \text{with} \quad |\tau_k\rangle = \sum_{j=0}^n \tau_{k,j} |j\rangle. \quad (48)$$

The key implication is that τ is now an effective size $(n + 1)$ matrix, and unlike ρ , does not have infinite dimensions. Now define

$$|T_k\rangle = \sum_{j=0}^n \tau_{k,j} |Gj\rangle. \quad (49)$$

From the form of our noise model, τ is a full rank matrix because it is a convex combination of a positive definite matrix, and a positive semi-definite matrix. The positive definite matrix arises from a truncation of the thermal state on the Fock states $|0\rangle, |G\rangle, \dots, |Gn\rangle$, and the positive semi-definite matrix arises from ρ_L . Since τ is a full rank matrix, it follows that

the spectral decomposition of ρ is

$$\begin{aligned} \rho = & \sum_{k=0}^n t_k |T_k\rangle\langle T_k| + \lambda_{\text{th}} \sum_{k=0}^{n-1} \sum_{j=1}^{g-1} \frac{e^{-\beta(Gk+j)}}{1 - e^{-\beta}} |Gk+j\rangle\langle Gk+j| \\ & + \lambda_{\text{th}} \sum_{k=Gn+1}^{\infty} \frac{e^{-\beta k}}{1 - e^{-\beta}} |k\rangle\langle k|. \end{aligned} \quad (50)$$

From the above spectral decomposition of ρ , it is clear that only the first summation term contributes to the state derivatives. This makes the effective dimension of the problem equal to the dimension of τ instead of that of ρ . Because of this reduction in the effective dimensionality of the problem, we can efficiently use algorithm 4 to evaluate upper and lower bounds to the HCRB for fixed u to benchmark parameter estimates for θ .

By optimising over u , we find exact values of the HCRB. In Fig. 8a we illustrate how the HCRB changes with different values of the noise parameters λ_{th} and β and the code parameters G and n . Notice that the effect of binomial codes in this application is limited at high thermalisation, where the state has a lot of thermal noise. The binomial codes are able to protect against errors resulting on the state due to low temperatures. To see this, we observe the behaviour of the HCRB in the region where it is minimised with respect to x . In Fig. 8b we illustrate the behaviour of the HCRB within $x = [0.68, 0.84]$ for different values of G . Notice that increasing the number of correctable amplitude damping errors by increasing the value of G improves the precision of the simultaneous estimate for x and ϕ . In Fig. 8c, we illustrate a similar improvement by increasing n , which is related to the number of correctable phase errors. This demonstrates that the error correcting codes can be used to improve simultaneous parameter estimates in the low error regime.

It is worth noting the performance of the simple bounds in theorem 1 for this application. For the specific choice of parameters $u = 0$, $x = -0.8$, $\phi = 0.7$, $n = 5$, $G = 3$, $\lambda_{\text{th}} = 0.2$, $\beta = 0.1$, we get the exact HCRB. In general, theorem 1 returns non-tight bounds to the HCRB for alternative values in the parameter space. This illustrates that if the tightness of bounds are crucial, then one should apply theorem 3.

V. CONCLUSIONS AND DISCUSSIONS

Quantum metrology promises practical near term quantum technologies. Experimental developments in sensing are demonstrating early theoretical results and advancements in estimation theory. On the theoretical front, one prominent limitation that remains is the estimation of multiple non-compatible observables. Specifically, the optimal strategy to define the fundamental limits to precision estimates and their attainability is not known. Efforts to estimate multiple non-compatible observables have largely been focused on approaching the fundamental quantum Cramér-Rao bound (QCRB). This has led to efforts to devise non trivial measurement schemes that approach the QCRB. An alternative approach is to focus on the tighter Holevo Cramér-Rao bound (HCRB), which is physically attainable. However, the HCRB is difficult to evaluate since it

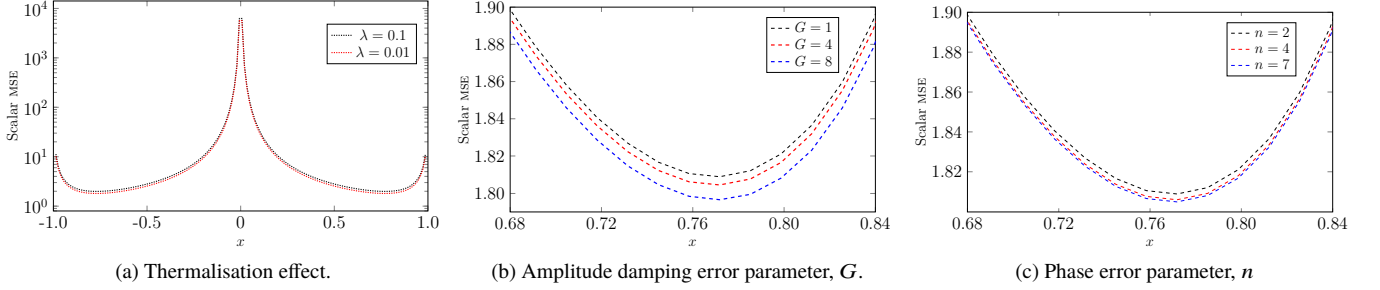


FIG. 8: Variation of HCRB with thermalisation and binomial code parameters. For the plot in Fig. 8a, we take $n = 2$, $\beta = 0.01$, and $G = 1$. In Fig. 8b, we see the effect of noise by varying the number of correctable amplitude damping errors of the binomial code, using $n = 2$, $\lambda_{\text{th}} = 0.01$, and $\beta = 1$. In Fig. 8c we illustrate the effect of increasing the correctable phase errors of the binomial code on the HCRB using $\lambda_{\text{th}} = 0.01$, $\beta = 1$, and $G = 1$.

involves a difficult optimisation over two observables. This has limited its application in quantum estimation theory.

In this paper, we have made significant progress in analytically solving the HCRB for two-parameter estimation problems, and providing bounds for larger number of parameters. In the two-parameter case, we reduce the complexity of the optimisation procedure to that of solving a set of linear equations, which can be easily solved using most numerical software packages. We also provide analytic expressions for the optimal positive operator valued measurements (POVM). Our results readily apply to a large range of physical applications. This will provide deeper insight into the role of quantum measurements in quantum sensing, and help continue the drive of realising quantum technologies.

We illustrate an application of our results by considering the estimation of a magnetic field using noisy multi-qubit probe states [74, 86]. A recent numerical study by Albarelli *et al.* demonstrated the necessity of using the HCRB over the QCRB, based on a violation of the weak commutation condition [44]. Here, we provide further insight into the role that the HCRB plays in quantum estimation theory. We provide conditions for when this bound is tighter than the SLD QCRB (or the Helstrom bound) and provide the corresponding optimal measurement observables.

A second application of our results explores how bosonic quantum error correction codes can improve noise resilience of parameter estimates. Bosonic codes are interesting because of their potential in reducing the number of physical systems required while having some robustness against errors. However, the infinite dimensionality of bosonic systems renders a brute force numerical approach to determining the HCRB intractable. Instead, through our analytical approach, we reduce this problem to a finite dimensional problem, and thereby evaluate the corresponding precision bounds efficiently.

There are several clear extensions of our work that can be readily addressed. The first would be to use the analytic expressions that we derive to provide further insight into more protocols in estimation theory. We hope that this will help to drive the wave for experimental validation. A second line of work would consider an extension of the Holevo bound to parameters with arbitrary choice of weight matrices. In this work

we have considered unit weight matrices, which was motivated through placing equal importance to each parameter. A more general weight matrix would provide a more general bound. A final line of work would consider the optimal implementation of the general measurements that were derived in this work. This would provide an immediate access to the tighter HCRB through experimental implementation.

VI. ACKNOWLEDGMENTS

J.S.S and P.K. acknowledge the support of EPSRC via the Quantum Communications Hub (EP/M013472/1). Y.O. and E.C. acknowledge the support of EPSRC through grant number EP/M024261/1.

Appendix A: Matrix calculus

We prove some elementary facts about matrix calculus that we use repeatedly in our analysis of the turning points of the Lagrangian functions that occur throughout the manuscript.

We begin by defining some notations. Since a complex matrix of size n is a map from \mathbb{C}^n to \mathbb{C}^n , we use $L(\mathbb{C}^n)$ to denote the set of all complex matrices of size n . Here, the notation $L(\mathbb{C}^n)$ reflects the fact that a matrix is a linear mapping that is an automorphism on \mathbb{C}^n . At times we are interested in matrices that are also Hermitian, which means that they are equal to their complex conjugates. In this scenario, we use \mathbb{H}_n to denote the set of all complex matrices that are also Hermitian. Clearly for instance, \mathbb{H}_n is a strict subset of $L(\mathbb{C}^n)$.

Now let $f : L(\mathbb{C}^n) \rightarrow \mathbb{C}$ denote a function that maps a complex matrix to a complex scalar. If $f(Y)$ is differentiable at Y in the direction H , we use

$$\nabla_{Y,H} f(Y) = \lim_{h \rightarrow 0} \frac{f(Y + hH) - f(Y)}{h} \quad (\text{A1})$$

to denote the Fréchet derivative of $f(Y)$ in the direction H . In the above formula, h is a real infinitesimal parameter. Properties of these Fréchet derivatives continues to be an active

area of research [87], and they have also been recently used in quantum information theory [88].

In this paper, we are interested in matrix functions that are either linear or quadratic in the matrix variable Y . This leads us to analyse the Fréchet derivatives given by the following lemma.

Lemma 4. *Let $Y, H \in L(\mathbb{C}^n)$. Then*

$$\nabla_{Y,H} \text{Tr}[AY] = \text{Tr}[AH] \quad (\text{A2})$$

$$\nabla_{Y,H} \text{Tr}[AY^\dagger] = \text{Tr}[AH^\dagger] \quad (\text{A3})$$

$$\nabla_{Y,H} \text{Tr}[YAY^\dagger] = \text{Tr}[AY^\dagger H + YAH^\dagger] \quad (\text{A4})$$

$$\nabla_{Y,H} \text{Tr}[Y^\dagger AY] = \text{Tr}[Y^\dagger AH + AYH^\dagger]. \quad (\text{A5})$$

Proof. The proof of the above results from direct application of the definition of the Fréchet derivative for the first two equations. For the last two equations, we also use the cyclic property of the trace. \square

We are often faced with the unconstrained minimisation of a quadratic form, and we show in the following lemma what the optimal solution to these optimisation problems are.

Lemma 5. *Let $A \in L(\mathbb{C}^n)$ and let ρ be a full rank matrix in \mathbb{H}_n . Then*

$$\min_{Y \in L(\mathbb{C}^n)} \left(\text{Tr}[Y\rho Y^\dagger] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger] \right) = -\text{Tr}[A^\dagger \rho^{-1} A] \quad (\text{A6})$$

$$\min_{Y \in L(\mathbb{C}^n)} \left(\text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger] \right) = -\text{Tr}[A \rho^{-1} A^\dagger], \quad (\text{A7})$$

with optimal solutions given by $Y = -A^\dagger \rho^{-1}$ and $Y = -\rho^{-1} A^\dagger$ respectively.

Proof. We first prove (A6) and (A7). The corresponding objective functions that are to be minimised are convex and differentiable, so it suffices to find when their Fréchet derivatives are equal to zero for any direction H . For this, we use Lemma 4, from which we find that we must have $\rho Y^\dagger + A = 0$ and $Y^\dagger \rho + A = 0$ respectively. Making use of the fact that ρ is invertible whenever it has full rank, we multiply both sides of the equations, and find that the optimal Y s are given by $Y = -A^\dagger \rho^{-1}$ and $Y = -\rho^{-1} A^\dagger$ respectively. Substituting this back into the objective functions gives the result. \square

Appendix B: Simple two-parameter bounds to the HCRB

We explicitly derive the HCRB for the two-parameter case. In the two-parameter setting, the HCRB with a weight matrix W is given by the optimisation problem

$$\begin{aligned} & \underset{X_1, X_2}{\text{minimize}} && \text{Tr}[W \text{Re}Z] + \|W \text{Im}Z\|_1, \\ & \text{subject to} && \text{Tr}[\rho X_j] = 0, \\ & && \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, \end{aligned} \quad (\text{B1})$$

where X_j are constrained to be Hermitian matrices in \mathbb{H}_N , and Z is a matrix given by

$$Z = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] \end{pmatrix}. \quad (\text{B2})$$

Note that W is always taken to be a positive definite matrix. For simplicity, we only consider the scenario where W is the identity matrix.

1. Reformulation of the optimisation problem

The optimisation problem (9) can be solved analytically primarily from our ability to rewrite the objective function as a quadratic function in the optimisation variables X_1 and X_2 . The method of Lagrange multipliers when applied to problems with quadratic objective functions and linear equality constraints is well-known to be exactly solvable, for example in theory of portfolio optimisation in finance [89]. A similar argument will allow us to solve (9) using this method.

We begin by showing why the objective function is quadratic. To see this, we first note that the diagonal terms of Z are positive numbers, because X_1 and X_2 are Hermitian and $X(\cdot)X^\dagger$ is a completely positive map. Second, the positivity of the diagonal entries of Z implies that

$$\text{ReTr}[Z] = \text{Tr}[Z] = \text{Tr}[X_1 \rho X_1^\dagger] + \text{Tr}[X_2 \rho X_2^\dagger].$$

Third, the positivity of the diagonal entries of Z implies that the trace norm of $\text{Im}Z$ can be explicitly evaluated. This is because the diagonal entries of $\text{Im}Z$ must be zero. Since X_1, X_2 and ρ are Hermitian matrices, it follows that

$$\text{Im}Z = \frac{1}{2i} \begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix},$$

where $w = \text{Tr}[\rho X_1 X_2] - \text{Tr}[X_2 X_1 \rho]$ is an imaginary number. The eigenvalues of $\text{Im}Z$ are therefore $\pm w/2$, which implies that the trace norm of $\text{Im}Z$ is $\max\{iw, -iw\}$. From this, we get

$$\text{ReTr}[Z] + iw = \text{Tr}[(X_1 + iX_2)\rho(X_1 + iX_2)^\dagger] \quad (\text{B3})$$

$$\text{ReTr}[Z] - iw = \text{Tr}[(X_1 - iX_2)\rho(X_1 - iX_2)^\dagger]. \quad (\text{B4})$$

Now let us make the substitution $Y = X_1 + iX_2$. In this scenario, we can rewrite the equality constraints in (B1) as

$$\begin{aligned} \text{Tr}[\rho Y] &= 0 \\ \text{Tr}[\partial_1 \rho Y] &= 1 \\ \text{Tr}[\partial_2 \rho Y] &= i. \end{aligned} \quad (\text{B5})$$

Hence the optimisation problem (9) can be written as

$$\begin{aligned}
& \underset{Y, t}{\text{minimize}} && t, \\
& \text{subject to} && \text{Tr}[Y\rho Y^\dagger] \leq t, \\
& && \text{Tr}[Y^\dagger \rho Y] \leq t, \\
& && \text{Tr}[\rho Y] = 0, \\
& && \text{Tr}[\partial_1 \rho Y] = 1, \\
& && \text{Tr}[\partial_2 \rho Y] = i.
\end{aligned} \tag{B6}$$

Note that the optimisation problem (B6) is a linear optimisation problem with convex quadratic and linear constraints. When the equality constraints are satisfied, the quadratic terms in the inequality constraints are non-negative, and by setting t to be arbitrarily large, we can see that the inequality constraints in (B6) can always be strictly satisfied. Since (B6) is also a convex optimisation problem because of its linear objective function and convex constraint functions, the Slater constraint qualification holds with respect to (B6). This implies that the first order Karush-Kuhn-Tucker (KKT) conditions suffices to determine the optimality of (B6).

2. Analysing the Lagrangian

The KKT conditions are stated in terms of the Lagrangian of (B6). The column vector of Lagrange multipliers corresponding to the equality constraints is

$$\mathbf{z} = (y_1, z_{1,1}, z_{2,1}, y_2, z_{1,2}, z_{2,2}). \tag{B7}$$

The Lagrangian of (B6) is

$$\begin{aligned}
\mathcal{L}(Y, t, u, v, \mathbf{z}) = & t + u\text{Tr}[Y\rho Y^\dagger] - ut + v\text{Tr}[Y^\dagger \rho Y] - vt \\
& + y_1\text{ReTr}[\rho Y] + y_2\text{ImTr}[\rho Y] \\
& + z_{1,1}(\text{ReTr}[\partial_1 \rho Y] - 1) + z_{1,2}\text{ImTr}[\partial_1 \rho Y] \\
& + z_{2,1}\text{ReTr}[\partial_2 \rho Y] + z_{2,2}(\text{ImTr}[\partial_2 \rho Y] - 1),
\end{aligned} \tag{B8}$$

where u, v are non-negative Lagrange multipliers corresponding to the inequality constraints.

There are four types of KKT conditions. First is the stationarity of the derivative of the Lagrangian with respect to the primal variables. Second is complementary slackness, which states that the product of the constraint functions [93] and their corresponding Lagrange multipliers is always zero. Third is the feasibility of the primal variables, and fourth is feasibility of the dual variables. If these KKT optimality conditions hold, then we can obtain the optimal solution and value of the corresponding optimisation problem.

Now we use the fact that

$$\text{ReTr}[\rho Y] = \frac{\text{Tr}[\rho Y] + \text{Tr}[\rho Y^\dagger]}{2} \tag{B9}$$

$$\text{ImTr}[\rho Y] = \frac{\text{Tr}[\rho Y] - \text{Tr}[\rho Y^\dagger]}{2i} \tag{B10}$$

$$\text{ReTr}[\partial_j \rho Y] = \frac{\text{Tr}[\partial_j \rho Y] + \text{Tr}[\partial_j \rho Y^\dagger]}{2} \tag{B11}$$

$$\text{ImTr}[\partial_j \rho Y] = \frac{\text{Tr}[\partial_j \rho Y] - \text{Tr}[\partial_j \rho Y^\dagger]}{2i}. \tag{B12}$$

Using this, it follows that

$$\begin{aligned}
\mathcal{L}(Y, t, u, v, \mathbf{z}) = & t(1 - u - v) - b^\top \mathbf{z} + u\text{Tr}[Y\rho Y^\dagger] \\
& + v\text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger],
\end{aligned} \tag{B13}$$

where b is the column vector $(0, 1, 0, 0, 0, 1)$ and

$$A = y_1 A_1 + z_{1,1} A_2 + z_{2,1} A_3 + y_2 A_4 + z_{1,2} A_5 + z_{2,2} A_6, \tag{B14}$$

where

$$A_1 = \frac{\rho}{2}, \quad A_2 = \frac{\partial_1 \rho}{2}, \quad A_3 = \frac{\partial_2 \rho}{2} \tag{B15}$$

and $\{A_4, A_5, A_6\} = -i\{A_1, A_2, A_3\}$.

Before we proceed to derive the Lagrange dual function, we note the following.

1. We prove that the optimal t must be strictly positive from the positive definiteness of ρ . From the positive definiteness of ρ , t is equal to zero if and only if Y is 0, but this would violate the feasibility constraints. Hence t cannot be equal to zero.
2. We prove that for optimal dual variables, we must have $u + v = 1$. If the dual variables u and v are such that $u + v < 1$, then the optimal t is 0 which is a contradiction. If $u + v > 1$, the optimal t is arbitrarily large which makes the Lagrangian unbounded from below, which is also a trivial scenario. This implies that the term in the Lagrangian $t(1 - u - v)$ for optimal values of the primal and dual variables must evaluate to zero. Hence for the optimal solution, we must have $u + v = 1$ and $t > 0$.
3. The KKT conditions require that the complementary slackness conditions hold for the inequality constraints in (B6). This means that

$$\begin{aligned}
u(\text{Tr}[Y\rho Y^\dagger] - t) &= 0 \\
v(\text{Tr}[Y^\dagger \rho Y] - t) &= 0
\end{aligned} \tag{B16}$$

If $\text{Tr}[Y\rho Y^\dagger] \neq \text{Tr}[Y^\dagger \rho Y]$, exactly one of the constraints corresponding to u and v must be tight, and complementary slackness implies that the optimal (u, v) must be either $(u, v) = (1, 0)$ or $(u, v) = (0, 1)$. This corresponds to the scenario where the QCRB is not equal to the HCRB. If $\|\text{Im}Z\|_1 = 0$, $(u, v) = (1, 0)$ and $(u, v) = (0, 1)$ do not

necessarily optimize the value of the Lagrange dual, and in general provide a lower bound to the Lagrange dual. However if $\text{Tr}[Y\rho Y^\dagger] = \text{Tr}[Y^\dagger\rho Y]$, then the ansatzes $(u, v) = (0, 1)$ and $(u, v) = (1, 0)$ will not yield tight bounds, because complementary slackness will not further constrain the optimal values of u and v .

3. Deriving the Lagrange dual functions

When $(u, v) = (1, 0)$, the Lagrangian evaluates to

$$\mathcal{L}(Y, t, 1, 0, z) = -\mathbf{b}^\top \mathbf{z} + \text{Tr}[Y\rho Y^\dagger] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger], \quad (\text{B17})$$

where $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$. Since ρ is full rank, ρ is invertible. Using Lemma 5, the above is minimised with respect to Y when $Y = -A^\dagger \rho^{-1}$ with optimal value $-\text{Tr}[A^\dagger \rho^{-1} A]$. In this scenario, the Lagrange dual function of (B6) evaluated with $(u, v) = (1, 0)$ is

$$g(1, 0, z) = -\text{Tr}[A^\dagger \rho^{-1} A] - \mathbf{b}^\top \mathbf{z}. \quad (\text{B18})$$

Similarly when $(u, v) = (0, 1)$, the Lagrangian evaluates to

$$\mathcal{L}(Y, t, 0, 1, z) = -\mathbf{b}^\top \mathbf{z} + \text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger], \quad (\text{B19})$$

and is minimised when $Y = -\rho^{-1} A^\dagger$ with an optimal value of $-\text{Tr}[A \rho^{-1} A^\dagger]$. In this scenario, the Lagrange dual function of (B6) evaluated with $(u, v) = (1, 0)$ is

$$g(0, 1, z) = -\text{Tr}[A \rho^{-1} A^\dagger] - \mathbf{b}^\top \mathbf{z}. \quad (\text{B20})$$

The Lagrange dual functions $g(1, 0, z)$ and $g(0, 1, z)$ can be rewritten in terms of the matrices Q_1 and Q_2 where in the Dirac bra-ket notation, we have

$$\begin{aligned} Q_1 &= \sum_{j,k=1,\dots,6} \text{Tr}[A_j^\dagger \rho^{-1} A_k] |j\rangle \langle k|, \\ Q_2 &= \sum_{j,k=1,\dots,6} \text{Tr}[A_j \rho^{-1} A_k^\dagger] |j\rangle \langle k|. \end{aligned} \quad (\text{B21})$$

Here $|j\rangle$ denotes a column vector and $\langle k|$ denotes a row vector. The Lagrange dual function that we consider are thus

$$\begin{aligned} g(1, 0, \mathbf{z}) &= -\mathbf{z}^\top Q_1 \mathbf{z} - b^\top \mathbf{z}, \\ g(0, 1, \mathbf{z}) &= -\mathbf{z}^\top Q_2 \mathbf{z} - b^\top \mathbf{z}. \end{aligned} \quad (\text{B22})$$

4. Upper and lower bounds

Using the fact that the Lagrange dual functions (B22) and the dual variables are real, lower bounds to the HCRB are given by

$$\max_{j=1,2} \max_{\mathbf{z} \in \mathbb{R}^6} (-\mathbf{z}^\top \text{Re}(Q_j) \mathbf{z} - \mathbf{b}^\top \mathbf{z}), \quad (\text{B23})$$

from which it follows that

$$2\text{Re}(Q_j) \mathbf{z} + \mathbf{b} = 0$$

is the correct optimality condition to consider. Thus, when $\text{Re}(Q_j)$ is full rank, the lower bounds to the HCRB can be written as

$$\max_{j=1,2} l_j, \quad \text{where} \quad l_j = \frac{1}{4} \mathbf{b}^\top \text{Re}(Q_j)^{-1} \mathbf{b} \quad (\text{B24})$$

Interestingly, when ρ is full rank, the matrices $\text{Re}(Q_j)$ are also full rank. We demonstrate this in the next subsection.

To obtain upper bounds to the HCRB, we appeal to the form of the primal problem is closely related to (B6), that has an objective function of

$$\max\{\text{Tr}[Y\rho Y^\dagger], \text{Tr}[Y^\dagger \rho Y]\}. \quad (\text{B25})$$

The upper bounds will be expressed in terms of the dual variables that optimise (B23), which we can write as $\mathbf{z}_1 = (z_{1,1}, \dots, z_{6,1})$ and $\mathbf{z}_2 = (z_{1,2}, \dots, z_{6,2})$ where

$$z_{a,j} = -\frac{1}{2}([\text{Re}(Q_j)^{-1}]_{a2} + [\text{Re}(Q_j)^{-1}]_{a6}). \quad (\text{B26})$$

and $[\text{Re}(Q_j)^{-1}]_{ab}$ denotes the matrix element in the a th row and b th column of $\text{Re}(Q_j)$.

Recall that when $(u, v) = (1, 0)$, the optimal solution to Y in minimising the Lagrangian is $-A^\dagger \rho^{-1}$. By choosing $A = A_1 z_{1,1} + \dots + A_6 z_{6,1}$, we find that the HCRB is thus upper bounded by

$$P_1 = \max\{\text{Tr}[A^\dagger \rho^{-1} A], \text{Tr}[\rho^{-2} A \rho A^\dagger]\}. \quad (\text{B27})$$

Then,

$$P_1 = \max\{l_1, m_1\}, \quad (\text{B28})$$

where

$$m_j = \sum_{a,b=1}^6 \text{Tr}[\rho^{-2} A_a \rho A_b^\dagger] z_{aj} z_{bj}. \quad (\text{B29})$$

When $(u, v) = (0, 1)$, the optimal solution to Y in minimising the Lagrangian is $-\rho^{-1} A^\dagger$. In this case the primal objective function is equal to

$$P_2 = \max\{l_2, m_2\}. \quad (\text{B30})$$

Hence the HCRB is at most

$$\min_{j=1,2} \{\max\{l_j, m_j\}\}. \quad (\text{B31})$$

This concludes the proof of theorem 1 for bounds on the two-parameter HCRB.

5. Full-rankness of Q

The analytic solution to the HCRB requires $\text{Re}(Q_j)$ to have full rank such that the solution can be determined. In this subsection, we demonstrate that the full-rankness of the probe state ρ entails the full-rankness of these matrices. Since the

regularity conditions of estimation theory require the state to be full-rank, our solution to the HCRB always exists.

Notice that the matrices Q_j defined in Eq. (B21) can be written

$$Q_1 = \begin{pmatrix} H & -iH \\ iH & H \end{pmatrix}, \quad Q_2 = \begin{pmatrix} H & iH \\ -iH & H \end{pmatrix}, \quad (\text{B32})$$

where H is Gram matrix defined as follows. We consider the Hilbert-Schmidt inner-product $\langle X, Y \rangle = \text{Tr}[X^\dagger Y]$. We define the operators

$$B_1 = \rho^{-1/2} A_1 = \rho^{-1/2} \rho / 2 = \rho^{1/2} / 2, \quad (\text{B33})$$

$$B_2 = \rho^{-1/2} A_2 = \rho^{-1/2} \delta_1 \rho / 2, \quad (\text{B34})$$

$$B_3 = \rho^{-1/2} A_3 = \rho^{-1/2} \delta_2 \rho / 2. \quad (\text{B35})$$

Then, we have that H is a Gram matrix with respect to this set of operators

$$H_{i,j} = \langle B_i, B_j \rangle. \quad (\text{B36})$$

As a Gram matrix, it is positive semi-definite. Furthermore, we know that H will be full-rank if and only if the set $\{B_1, B_2, B_3\}$ is linearly independent. We note that A_1 cannot be written as a sum of A_2 and A_3 (since A_1 has nonzero trace whereas A_2 and A_3 are traceless). Also by a trace-argument, if $\{A_1, A_2, A_3\}$ are linearly dependent, we must have that A_2 is proportional to A_3 . But if A_2 and A_3 are proportional, then it is really a one-parameter problem and not a two-parameter problem. Hence, $\{A_1, A_2, A_3\}$ are linearly independent. If we assume ρ is full-rank, then $\{B_1, B_2, B_3\} = \rho^{-1/2} \{A_1, A_2, A_3\}$ is also a linearly independent set. So full-rankness of ρ entails full-rankness of H .

We are now interested in the real part of the Q_j matrices, looking for solutions of $2\text{Re}(Q_j)z + b = 0$. Considering $j = 2$

$$\text{Re}(Q_2) = \begin{pmatrix} \text{Re}(H) & -\text{Im}(H) \\ \text{Im}(H) & \text{Re}(H) \end{pmatrix}, \quad (\text{B37})$$

By performing elementary row operations by taking a linear combination of rows, followed by elementary column operators by taking a linear combination of columns, we get

$$\text{Re}(Q_2) \rightarrow \begin{pmatrix} H & iH \\ \text{Im}(H) & \text{Re}(H) \end{pmatrix} \rightarrow \begin{pmatrix} H & 0 \\ \text{Im}(H) & H^* \end{pmatrix}, \quad (\text{B38})$$

where we used $\text{Re}(H) + i\text{Im}(H) = H$. Since both rows are linearly independent, $\text{Re}(Q_j)$ is also always full-rank. Therefore, we have that if the state is full-rank, then so too is the matrix $\text{Re}(Q_j)$.

Appendix C: Lower bound in the multi-parameter setting

By restricting ourselves to the identity weight matrix, recall that the HCRB is the optimal value of the following optimisation

problem over the Hermitian matrices X_j in \mathbb{H}_N given by

$$\begin{aligned} & \text{minimize}_{X_1, \dots, X_d} \quad \text{Tr}[\text{Re}Z] + \|\text{Im}Z\|_1, \\ & \text{subject to} \quad \text{Tr}[\rho X_j] = 0, \\ & \quad \quad \quad \text{Tr}[\partial_j \rho X_k] = \delta_{jk}. \end{aligned} \quad (\text{C1})$$

For $j, k = \{1, \dots, d\}$, let

$$w_{j,k} = \text{Tr}[\rho X_j X_k] - \text{Tr}[\rho X_k X_j]. \quad (\text{C2})$$

1. Deriving a lower bound for the objective function

In general for a d -parameter estimation problem, we have

$$\text{Tr}[\text{Re}Z] = \sum_{j=1}^d \text{Tr}[X_j \rho X_j^\dagger] \quad (\text{C3})$$

$$\text{Im}Z = \frac{1}{2i} \sum_{1 \leq j < k \leq d} w_{j,k} (|j\rangle\langle k| - |k\rangle\langle j|). \quad (\text{C4})$$

Note that $\text{Im}Z$ is always a skew-Hermitian matrix. For example, when $d = 3$, we have

$$\text{Im}Z = \begin{pmatrix} 0 & w_{1,2} & w_{1,3} \\ -w_{1,2} & 0 & w_{2,3} \\ -w_{1,3} & -w_{2,3} & 0 \end{pmatrix}. \quad (\text{C5})$$

We use several ansatz's to obtain a lower bound for the trace norm of $\text{Im}Z$ utilising the fact that

$$\|\text{Im}Z\|_1 = \max\{\text{Tr}[U \text{Im}Z] : -\mathbb{1}_d \leq U \leq \mathbb{1}_d\}. \quad (\text{C6})$$

Here, the inequality for the matrix indicates Loewner ordering of matrices. Instead of optimising over all possible U , we can optimise U over the finite set

$$\mathcal{U} = \left\{ \sum_{j=0}^{d-1} |i\rangle\langle i \oplus s| (-1)^{x_i} : x_i \in \{0, 1\}, s = 1, 2, \dots, d-1 \right\}, \quad (\text{C7})$$

where \oplus denotes addition modulo d over the ring \mathbb{Z}_d , s denotes the shift, and $\{|j\rangle : j = 0, \dots, d-1\}$ is an orthonormal basis. Clearly, the singular values of the matrices in \mathcal{U} are all equal to 1, and hence $\|\mathcal{U}\|_\infty \leq 1$, where $\|\cdot\|_\infty$ denotes the Schatten infinity norm. Hence we can use \mathcal{U} to get the lower bound

$$\|\text{Im}Z\|_1 \geq \max\{\text{Tr}[\text{Im}Z U] : U \in \mathcal{U}\}, \quad (\text{C8})$$

which implies that for every $s = \{1, \dots, \lfloor (d-1)/2 \rfloor\}$ we have

$$\|\text{Im}Z\|_1 \geq \frac{1}{2} \sum_{(a,b) \in E_{d,s}} |w_{a,b}| \quad (\text{C9})$$

where

$$E_{d,s} = \{(j+1, (j \oplus s) + 1) : j = 0, \dots, d-1\} \quad (\text{C10})$$

and each $|w_{a,b}|$ is the maximum of

$$i\text{Tr}[X_b\rho X_a^\dagger] - i\text{Tr}[X_a\rho X_b^\dagger]$$

and

$$-i\text{Tr}[X_b\rho X_a^\dagger] + i\text{Tr}[X_a\rho X_b^\dagger].$$

2. Recasting the optimisation problem

Now let us observe that $\text{Tr}[\text{Re}Z] + \sum_{(a,b) \in E_{d,s}} |w_{a,b}|/2$ can be written as a quadratic form in the optimisation variables X_1, \dots, X_d . To see this, note for instance that

$$\begin{aligned} & \text{Tr}[\text{Re}Z] + \frac{i}{2} \sum_{(a,b) \in E_{d,1}} w_{a,b} \\ &= \sum_{j=1}^d \text{Tr}[X_j\rho X_j^\dagger] + \frac{(i\text{Tr}[X_2\rho X_1^\dagger] + \dots + i\text{Tr}[X_1\rho X_d^\dagger])}{2} \\ & \quad - \frac{(i\text{Tr}[X_1\rho X_2^\dagger] + \dots + i\text{Tr}[X_d\rho X_1^\dagger])}{2} \\ &= \frac{1}{2} \sum_{(a,b) \in E_{d,s}} \text{Tr}[(X_a + iX_b)\rho(X_a + iX_b)^\dagger], \end{aligned} \quad (\text{C11})$$

The last equality in Eq. (C11) arises because every j can be a component of (a, b) in exactly two ways, as the tuples (a, b) correspond to edges in a cycle graph. To see how this works explicitly in the $d = 3$ and $s = 1$ scenario, note that

$$\begin{aligned} & \text{Tr}[\text{Re}Z] + \frac{i}{2} \sum_{(a,b) \in E_{d,1}} w_{a,b} \\ &= \text{Tr}[X_1\rho X_1^\dagger] + \text{Tr}[X_2\rho X_2^\dagger] + \text{Tr}[X_3\rho X_3^\dagger] \\ & \quad + \frac{i}{2}(\text{Tr}[X_2\rho X_1^\dagger] + \text{Tr}[X_3\rho X_2^\dagger] + \text{Tr}[X_1\rho X_3^\dagger]) \\ & \quad - \frac{i}{2}(\text{Tr}[X_1\rho X_2^\dagger] + \text{Tr}[X_2\rho X_3^\dagger] + \text{Tr}[X_3\rho X_1^\dagger]) \\ &= \frac{1}{2}(\text{Tr}[(X_1 + iX_2)\rho(X_1 + iX_2)^\dagger] \\ & \quad + \frac{1}{2}\text{Tr}[(X_2 + iX_3)\rho(X_2 + iX_3)^\dagger]) \\ & \quad + \frac{1}{2}\text{Tr}[(X_3 + iX_1)\rho(X_3 + iX_1)^\dagger]). \end{aligned} \quad (\text{C12})$$

Now, given a binary vector $\alpha = (\alpha_1, \dots, \alpha_d)$, let us define

$$V_\alpha = \frac{1}{2} \sum_{(a,b) \in E_{d,s}} \text{Tr}[(X_a + (-1)^{\alpha_a} iX_b)\rho(X_a + (-1)^{\alpha_a} iX_b)^\dagger]. \quad (\text{C13})$$

Then we can rewrite (C1) as an optimisation over Hermitian matrices X_1, \dots, X_d where

$$\begin{aligned} & \text{minimize}_{X_1, \dots, X_d} \quad \max_{\alpha \in \{0,1\}^d} V_\alpha, \\ & \text{subject to} \quad \text{Tr}[\rho X_j] = 0, \\ & \quad \text{Tr}[\partial_j \rho X_k] = \delta_{jk}. \end{aligned} \quad (\text{C14})$$

We can rewrite with an introduction of an auxillary variable $\alpha \in \mathbb{R}$ so that (C14) is equivalent to

$$\begin{aligned} & \text{minimize}_{X_1, \dots, X_d, t} \quad t, \\ & \text{subject to} \quad \text{Tr}[\rho X_j] = 0, \\ & \quad \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, \\ & \quad V_\alpha \leq t, \\ & \quad \alpha \in \{0, 1\}^d. \end{aligned} \quad (\text{C15})$$

This minimisation problem can be numerically checked for consistency with the optimisation in Eq. (C1).

3. Diagonalising the quadratic forms

At this point, we want to define new variables Y_1, \dots, Y_d so that we can for example write

$$V_\alpha = \frac{1}{2} \sum_{j=1}^d (\delta_{0,\alpha_j} \text{Tr}[Y_j \rho Y_j^\dagger] + \delta_{1,\alpha_j} \text{Tr}[Y_j^\dagger \rho Y_j]). \quad (\text{C16})$$

In general, we define the variables Y_1, \dots, Y_d so that they depend linearly on the variables X_1, \dots, X_d according to the system of linear equations

$$\mathbf{Y} = \mathbf{S}\mathbf{X}, \quad (\text{C17})$$

where

$$\mathbf{S} = \begin{pmatrix} S_{1,1}\mathbf{1} & \dots & S_{1,d}\mathbf{1} \\ \vdots & & \vdots \\ S_{d,1}\mathbf{1} & \dots & S_{d,d}\mathbf{1} \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_d \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}. \quad (\text{C18})$$

It is important to also express the variables X_1, \dots, X_d in terms of Y_1, \dots, Y_d , and this means that we need the matrix

$$S = \begin{pmatrix} S_{1,1} & \dots & S_{1,d} \\ \vdots & & \vdots \\ S_{d,1} & \dots & S_{d,d} \end{pmatrix} \quad (\text{C19})$$

to be invertible. When S is full rank, $T = S^{-1}$ exists. One possible choice of S is where $S_{j,j} = 1$ and $S_{j,j+1} = i$ for all $j = 1, \dots, d-1$, and $S_{d,d} = 1, S_{d,1} = i$, and all other matrix elements of S are zero. For instance, when $d = 3$, we use

$$S = \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ i & 0 & 1 \end{pmatrix}. \quad (\text{C20})$$

However such a choice of S need not have be full rank. It is easy to see that S is full rank whenever its dimension is not a multiple of 4.

Proposition 6. Let d be a positive integer, that is at least 2 and is not a multiple of 4. Then S has full rank.

Proof. To prove this, we use pairwise orthogonality of the vectors $(1, i), (1, -i), (1, -1)$ and $(1, 1)$. By Gaussian elimination and the properties of i , we can easily arrive at the result. \square

When d is not a multiple of 4, we denote

$$T = \begin{pmatrix} T_{1,1} & \dots & T_{1,d} \\ \vdots & & \vdots \\ T_{d,1} & \dots & T_{d,d} \end{pmatrix}, \quad (\text{C21})$$

and it follows that for every $j = 1, \dots, d$, we have

$$X_k = \sum_{\ell=1}^d T_{k,\ell} Y_\ell. \quad (\text{C22})$$

From (C22), assuming the Hermiticity of X_k , we can recast the equality constraints in (C15) as

$$c_{0,k}(\mathbf{Y}) = \frac{1}{2} \sum_{\ell=1}^d \left(T_{k,\ell} \text{Tr}[\rho Y_\ell] + T_{k,\ell}^* \text{Tr}[\rho Y_\ell^\dagger] \right) = 0, \quad (\text{C23})$$

$$c_{j,k}(\mathbf{Y}) = \frac{1}{2} \sum_{\ell=1}^d \left(T_{k,\ell} \text{Tr}[\rho_j Y_\ell] + T_{k,\ell}^* \text{Tr}[\rho_j Y_\ell^\dagger] \right) - \delta_{j,k} = 0. \quad (\text{C24})$$

By choosing our explicit form of the matrix A strategically, we can write the quadratic inequality constraints with V_t s in the form as given by (C16).

Since the variables Y_ℓ are non-Hermitian in general, we need to impose additional constraints, namely the fact that the corresponding X_k are Hermitian. The Hermiticity of X_k implies from (C22) that

$$\sum_{\ell=1}^d \left(T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger \right) = 0. \quad (\text{C25})$$

The left side of (C25) is in general an antihermitian matrix, and to make it Hermitian, we multiply both sides by i to get

$$H_k(\mathbf{Y}) = i \sum_{\ell=1}^d \left(T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger \right) = 0. \quad (\text{C26})$$

With all these constraints, we recast the optimisation problem (C15) as the following optimisation problem.

$$\begin{aligned} & \text{minimize} && t, \\ & Y_1, \dots, Y_d, t \\ & \text{subject to} && c_{0,k}(\mathbf{Y}) = 0, \\ & && c_{j,k}(\mathbf{Y}) = 0, \\ & && \frac{1}{2} \sum_{\ell=1}^d \left(\delta_{0,\alpha_\ell} \text{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\alpha_\ell} \text{Tr}[Y_\ell^\dagger \rho Y_\ell] \right) \leq t, \\ & && H_k(\mathbf{Y}) = 0. \end{aligned} \quad (\text{C27})$$

4. Analysis on the Lagrangian

Here we consider the constraints in (C27) over $j, k = 1, \dots, d$ and $\alpha_1, \dots, \alpha_d = 0, 1$, which gives us a total of $d(d+1)$ regular equality constraints, d matrix equality constraints, and 2^d regular inequality constraints. The Lagrangian corresponding to (C27) can then be written as

$$\begin{aligned} L_d = & t + \sum_{j=0}^d \sum_{k=1}^d z_{j,k} c_{j,k}(\mathbf{Y}) + \sum_{k=1}^d \text{Tr}[\xi_k H_k(\mathbf{Y})] \\ & + \frac{1}{2} \sum_{\alpha \in \{0,1\}^d} v_\alpha \sum_{\ell=1}^d \left(\delta_{0,\alpha_\ell} \text{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\alpha_\ell} \text{Tr}[Y_\ell^\dagger \rho Y_\ell] \right) \\ & - \sum_{\alpha \in \{0,1\}^d} v_\alpha t. \end{aligned} \quad (\text{C28})$$

Here, the Lagrange multipliers $z_{j,k}$ are real numbers while the Lagrange multipliers v_α are non-negative numbers. The Lagrange multipliers ξ_k are Hermitian matrices in \mathbb{H}_N . Note that the multiparameter Lagrangian is a quadratic form in \mathbf{Y} , and as such, can be minimised using Lemma 5. Before for we do so, we consider the minimisation of the Lagrangian with respect to the primal variable t .

If the Lagrangian multiplier v_α do not all sum to one, by picking t to either approach positive or negative infinity, the Lagrangian L_d becomes unbounded. Hence the optimal multipliers v_α must sum to one. By picking a discrete set of values of v_α where v_α is equal to zero to all but one value of α , and maximising the Lagrange dual function for each of these cases, we can obtain our lower bound to the multi-parameters HCRB.

Hence without loss of generality, there is some value of the binary vector α for which the effective Lagrangian that we need to consider is

$$\begin{aligned} L_{d,\alpha} = & \sum_{j=0}^d \sum_{k=1}^d z_{j,k} c_{j,k}(\mathbf{Y}) + \sum_{k=1}^d \text{Tr}[\xi_k H_k(\mathbf{Y})] \\ & + \frac{1}{2} \sum_{\ell=1}^d \left(\delta_{0,\alpha_\ell} \text{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\alpha_\ell} \text{Tr}[Y_\ell^\dagger \rho Y_\ell] \right). \end{aligned} \quad (\text{C29})$$

Now let $\rho_0 = \rho$, $\rho_j = \partial_j \rho$ for $j = 1, 2$, and let

$$\Gamma_\ell = \sum_{k=1}^d T_{k,\ell} \left(\sum_{j=0}^d z_{j,k} \rho_j + i \xi_k \right). \quad (\text{C30})$$

Then we can rewrite the terms on the first line on the right side of (C29) as

$$\begin{aligned} L_{d,\alpha} = & - \sum_{j=1}^d z_{j,j} + \frac{1}{2} \sum_{\ell=1}^d \left(\text{Tr}[\Gamma_\ell Y_\ell] + \text{Tr}[\Gamma_\ell^\dagger Y_\ell^\dagger] \right. \\ & \left. + \delta_{0,\alpha_\ell} \text{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\alpha_\ell} \text{Tr}[Y_\ell^\dagger \rho Y_\ell] \right). \end{aligned} \quad (\text{C31})$$

Then, given that ρ is a Hermitian full rank matrix, we can use

Lemma 5 to get the corresponding Lagrange dual to be

$$g_\alpha = \min_{\mathbf{Y}} \mathcal{L}_{d,\alpha}$$

$$= - \sum_{j=1}^d z_{j,j} - \sum_{\ell=1}^d \frac{\delta_{0,\alpha_\ell} \text{Tr}[\Gamma_\ell \rho^{-1} \Gamma_\ell^\dagger] + \delta_{1,\alpha_\ell} \text{Tr}[\Gamma_\ell^\dagger \rho^{-1} \Gamma_\ell]}{2}. \quad (\text{C32})$$

Our lower bound to the HCRB is thus

$$\max_{\alpha \in \{0,1\}^d} \max\{g_\alpha : z_{j,k} \in \mathbb{R}, \xi_k \in \mathbb{H}_D\}, \quad (\text{C33})$$

where $j = 0, \dots, d, k = 1, \dots, d$. Any feasible value of g_α yields a lower bound to the HCRB.

Appendix D: Minimising the Lagrangian

In this appendix, we extend our formalism to account for arbitrary values of the Lagrange dual variable u . When $u + v = 1$, we minimise the Lagrangian which we recall has the form

$$\mathcal{L}(Y, u, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + u \text{Tr}[Y \rho Y^\dagger] + (1-u) \text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger], \quad (\text{D1})$$

where

$$A = (z_1 \rho + z_2 \rho_1 + z_3 \rho_2 - i z_4 \rho - i z_5 \rho_1 - i z_6 \rho_2)/2. \quad (\text{D2})$$

Recall the notation $\rho_j = \partial_j \rho, j \in \{1, 2\}$. The Frechet derivative of the Lagrangian in the matrix direction H is given by

$$\nabla_Y(\mathcal{L}, H) = \lim_{h \rightarrow 0} \frac{\mathcal{L}(Y + hH, u, \mathbf{z}) - \mathcal{L}(Y, u, \mathbf{z})}{h}. \quad (\text{D3})$$

Lemma 7. Let \mathcal{L} be the Lagrangian as defined in (D1) and A be as given in (D2). Suppose that

$$uY\rho + (1-u)\rho Y + A^\dagger = 0. \quad (\text{D4})$$

Then $\nabla_Y(\mathcal{L}, H) = 0$.

Proof. Notice that

$$\nabla_Y(\mathcal{L}, H) = u(\text{Tr}[H\rho Y^\dagger] + \text{Tr}[Y\rho H^\dagger]) + (1-u)(\text{Tr}[H^\dagger \rho Y] + \text{Tr}[Y^\dagger \rho H]) + \text{Tr}[AH] + \text{Tr}[A^\dagger H^\dagger]. \quad (\text{D5})$$

Now we use the cyclic property of the trace to write

$$\begin{aligned} \nabla_Y(\mathcal{L}, H) &= u(\text{Tr}[\rho Y^\dagger H] + \text{Tr}[Y\rho H^\dagger]) + (1-u)(\text{Tr}[\rho Y H^\dagger] + \text{Tr}[Y^\dagger \rho H]) + \text{Tr}[AH] + \text{Tr}[A^\dagger H^\dagger] \\ &= \text{Tr}[B^\dagger H] + \text{Tr}[BH^\dagger], \end{aligned} \quad (\text{D6})$$

where $B = uY\rho + (1-u)\rho Y + A^\dagger$ and $B^\dagger = u\rho Y^\dagger + (1-u)Y^\dagger \rho + A$. Since $B = 0$ by the assumption of our lemma, we must have $B^\dagger = 0$, and it follows that $\nabla_Y(\mathcal{L}, H) = 0$. \square

Notice that Eq. (D4) is a Sylvester equation, and solving it is a standard procedure, where a variant of the Bartel-Stewart algorithm can apply. When ρ is a full-rank matrix, we can solve this equation analytically. Let ρ have the spectral decomposition $\rho = \sum_j p_j |e_j\rangle \langle e_j|$ where $|e_j\rangle$ are normalised eigenvectors of ρ . In this case, (D4) is equivalent to

$$u \sum_{j,k} \langle e_j | Y | e_k \rangle |e_j\rangle \langle e_k| p_k + (1-u) \sum_{j,k} p_j |e_j\rangle \langle e_k| \langle e_j | Y | e_k \rangle + \sum_{j,k} |e_j\rangle \langle e_k| \langle e_j | A^\dagger | e_k \rangle = 0. \quad (\text{D7})$$

Simplifying this we get

$$\sum_{j,k} \langle e_j | Y | e_k \rangle (u p_k + (1-u) p_j) |e_j\rangle \langle e_k| = - \sum_{j,k} |e_j\rangle \langle e_k| \langle e_j | A^\dagger | e_k \rangle, \quad (\text{D8})$$

from which it follows that

$$\langle e_j | Y | e_k \rangle = -(u p_k + (1-u) p_j)^{-1} \langle e_j | A^\dagger | e_k \rangle. \quad (\text{D9})$$

The following lemma then follows.

Lemma 8. *The Y that minimises the Lagrangian is given by*

$$Y = - \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | A^\dagger | e_k \rangle |e_j\rangle \langle e_k|. \quad (\text{D10})$$

Crucially, we can write the Y that minimises the Lagrangian as a linear combination of the Lagrange multipliers \mathbf{z} .

Lemma 9. *The Y that minimises the Lagrangian is given by*

$$Y = z_1 \gamma_1 + z_2 \gamma_2 + z_3 \gamma_3 + iz_4 \gamma_1 + iz_5 \gamma_2 + iz_6 \gamma_3. \quad (\text{D11})$$

where the complex matrices

$$\gamma_1 = -\mathbf{1}_d/2, \quad (\text{D12})$$

$$\gamma_2 = - \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho_1/2 | e_k \rangle |e_j\rangle \langle e_k|, \quad (\text{D13})$$

$$\gamma_3 = - \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho_2/2 | e_k \rangle |e_j\rangle \langle e_k|. \quad (\text{D14})$$

Proof. Recall the definition of A in Eq. (D2). Then, through Lemma 8, we find

$$\begin{aligned} Y = & - (z_1 + iz_4) \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho/2 | e_k \rangle |e_j\rangle \langle e_k| \\ & - (z_2 + iz_5) \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho_1/2 | e_k \rangle |e_j\rangle \langle e_k| \\ & - (z_3 + iz_6) \sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho_2/2 | e_k \rangle |e_j\rangle \langle e_k|. \end{aligned} \quad (\text{D15})$$

Now we can make the simplification

$$\sum_{j,k} (up_k + (1-u)p_j)^{-1} \langle e_j | \rho/2 | e_k \rangle |e_j\rangle \langle e_k| = \sum_j (up_j + (1-u)p_j)^{-1} p_j/2 |e_j\rangle \langle e_j| = \mathbf{1}_d/2. \quad (\text{D16})$$

Hence, the result follows. \square

By substituting the optimal value of Y back into the Lagrangian, we find that Lagrangian is a quadratic in \mathbf{z} . Namely, we have the following.

Lemma 10. *For fixed u such that $0 < u < 1$, and where $\mathbf{z} = (z_1, \dots, z_6) \in \mathbb{R}^6$, the Lagrange dual of our Lagrangian is*

$$g(u, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + \mathbf{z}^\top \mathbf{Q} \mathbf{z}, \quad (\text{D17})$$

where

$$\mathbf{Q} = u\text{Re}(\mathbf{Q}_1) + (1-u)\text{Re}(\mathbf{Q}_2) + \text{Re}(\mathbf{Q}_3), \quad (\text{D18})$$

and

$$\mathbf{Q}_1 = \begin{pmatrix} G_1 & -iG_1 \\ iG_1 & G_1 \end{pmatrix}, \quad \mathbf{Q}_2 = \begin{pmatrix} G_2 & iG_2 \\ -iG_2 & G_2 \end{pmatrix}, \quad \mathbf{Q}_3 = \frac{1}{2} \begin{pmatrix} G_3 + G_3^* & i(G_3 - G_3^*) \\ -i(G_3 - G_3^*) & G_3 + G_3^* \end{pmatrix}, \quad (\text{D19})$$

and

$$G_1 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2 \rho \gamma_2^\dagger] & \text{Tr}[\gamma_2 \rho \gamma_3^\dagger] \\ 0 & \text{Tr}[\gamma_3 \rho \gamma_2^\dagger] & \text{Tr}[\gamma_3 \rho \gamma_3^\dagger] \end{pmatrix}, \quad G_2 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2^\dagger \rho \gamma_2] & \text{Tr}[\gamma_2^\dagger \rho \gamma_3] \\ 0 & \text{Tr}[\gamma_3^\dagger \rho \gamma_2] & \text{Tr}[\gamma_3^\dagger \rho \gamma_3] \end{pmatrix}, \quad G_3 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\rho_1 \gamma_2] & \text{Tr}[\rho_1 \gamma_3] \\ 0 & \text{Tr}[\rho_2 \gamma_2] & \text{Tr}[\rho_2 \gamma_3] \end{pmatrix} \quad (\text{D20})$$

Proof. The Lagrange dual is given by substituting the optimal solution for Y in the Lagrangian minimization. Recall the definition of the Lagrangian in Eq. (D1), then the first term to evaluate is

$$\text{Tr}[Y\rho Y^\dagger] = \text{Tr}[(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3)\rho(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)], \quad (\text{D21})$$

where we used the optimal solution for Y as given in Lemma 9. Writing this in matrix form, we have

$$\text{Tr}[Y\rho Y^\dagger] = \mathbf{z}^\top \begin{pmatrix} G_1 & -iG_1 \\ iG_1 & G_1 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_1 \mathbf{z}, \quad (\text{D22})$$

where \mathbf{z} is the column vector of Lagrange multipliers and the block matrix

$$G_1 = \begin{pmatrix} \text{Tr}[\gamma_1\rho\gamma_1^\dagger] & \text{Tr}[\gamma_1\rho\gamma_2^\dagger] & \text{Tr}[\gamma_1\rho\gamma_3^\dagger] \\ \text{Tr}[\gamma_2\rho\gamma_1^\dagger] & \text{Tr}[\gamma_2\rho\gamma_2^\dagger] & \text{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ \text{Tr}[\gamma_3\rho\gamma_1^\dagger] & \text{Tr}[\gamma_3\rho\gamma_2^\dagger] & \text{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix} = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2\rho\gamma_2^\dagger] & \text{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ 0 & \text{Tr}[\gamma_3\rho\gamma_2^\dagger] & \text{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix}. \quad (\text{D23})$$

Here, we have used the definitions for γ_j in Lemma 9 and the traceless property of the state derivatives ρ_1 and ρ_2 :

$$\text{Tr}[\rho\gamma_s] = \text{Tr}[\rho\gamma_s^\dagger] = \sum_j p_j((1-u)p_j + up_j)^{-1} \langle e_j | \rho_{s-1} / 2 | e_j \rangle = \sum_j \langle e_j | \rho_{s-1} | e_j \rangle / 2 = 0, \quad (\text{D24})$$

for $s \in \{2, 3\}$. Similarly, we determine the second term in the Lagrangian

$$\text{Tr}[Y^\dagger \rho Y] = \text{Tr}[(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)\rho(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3)]. \quad (\text{D25})$$

In matrix form, this can similarly be written as

$$\text{Tr}[Y^\dagger \rho Y] = \mathbf{z}^\top \begin{pmatrix} G_2 & iG_2 \\ -iG_2 & G_2 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_2 \mathbf{z} \quad (\text{D26})$$

where

$$G_2 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2^\dagger \rho \gamma_2] & \text{Tr}[\gamma_2^\dagger \rho \gamma_3] \\ 0 & \text{Tr}[\gamma_3^\dagger \rho \gamma_2] & \text{Tr}[\gamma_3^\dagger \rho \gamma_3] \end{pmatrix} \quad (\text{D27})$$

Finally, we substitute the optimal solution for Y into the last two terms of the Lagrangian

$$\begin{aligned} \text{Tr}[AY + A^\dagger Y^\dagger] &= \frac{1}{2} \text{Tr}[(z_1\rho + z_2\rho_1 + z_3\rho_2 - iz_4\rho - iz_5\rho_1 - iz_6\rho_2)(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3) \\ &\quad + (z_1\rho + z_2\rho_1 + z_3\rho_2 + iz_4\rho + iz_5\rho_1 + iz_6\rho_2)(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)], \end{aligned} \quad (\text{D28})$$

where we used the Hermiticity of the state derivatives. Hence

$$\text{Tr}[AY + A^\dagger Y^\dagger] = \frac{1}{2} \mathbf{z}^\top \begin{pmatrix} G_3 & iG_3 \\ -iG_3 & G_3 \end{pmatrix} \mathbf{z} + \frac{1}{2} \mathbf{z}^\top \begin{pmatrix} G_4 & -iG_4 \\ iG_4 & G_4 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_3 \mathbf{z}, \quad (\text{D29})$$

where

$$G_3 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\rho_1\gamma_2] & \text{Tr}[\rho_1\gamma_3] \\ 0 & \text{Tr}[\rho_2\gamma_2] & \text{Tr}[\rho_2\gamma_3] \end{pmatrix}, \quad G_4 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\rho_1\gamma_2^\dagger] & \text{Tr}[\rho_1\gamma_3^\dagger] \\ 0 & \text{Tr}[\rho_2\gamma_2^\dagger] & \text{Tr}[\rho_2\gamma_3^\dagger] \end{pmatrix}. \quad (\text{D30})$$

Notice that $G_4 = G_3^*$. Since the Lagrange dual must be real, and the dual variables u and \mathbf{z} must be real, we can take the real part of the matrices Q_1 , Q_2 and Q_3 to complete the proof. \square

Because our optimization problem is convex, we are promised that $L(Y, u, \mathbf{z})$ will be a concave function in both

u and \mathbf{z} . This implies that for fixed u , $L(Y, u, \mathbf{z})$ is concave in \mathbf{z} , which implies that Q is negative definite. From the stationary point of $g(u, \mathbf{z})$ with respect to \mathbf{z} , we find that the optimal solution to \mathbf{z} is given by the solution to the linear equation

$$2Q\mathbf{z} = \mathbf{b}. \quad (\text{D31})$$

If Q is full rank, then we reach the optimal values for the Lagrange multipliers

$$\mathbf{z} = \frac{1}{2}Q^{-1}\mathbf{b}. \quad (\text{D32})$$

Substituting this into the solution for Y in Eq. (D11) provides an ansatz that can saturate the HCRB, which is upper and lower bounded by

$$\begin{aligned} \mathcal{U}_u &= \max \{ \mathbf{z}^\top Q_1 \mathbf{z}, \mathbf{z}^\top Q_2 \mathbf{z} \}, \\ \mathcal{L}_u &= -\mathbf{b}^\top \mathbf{z} + \mathbf{z}^\top Q \mathbf{z} \end{aligned} \quad (\text{D33})$$

respectively. For fixed u , our ansatz for Y gives a tight bound when these two bounds are equivalent. If such a solution exists, we can optimise over the dual variable u , to find the optimal value. This can be determined numerically for any application in $O(\text{polylog}(1/\epsilon))$ time, where ϵ is the duality gap. Alternatively, we can find the optimal u by looking solely at the lower bound to the HCRB.

$$\begin{aligned} \mathcal{L}_u &= -\frac{1}{2}\mathbf{b}^\top Q^{-1}\mathbf{b} + \frac{1}{4}\mathbf{b}^\top (Q^{-1})^\top Q Q^{-1}\mathbf{b} \\ &= -\frac{1}{4}\mathbf{b}^\top Q^{-1}\mathbf{b}, \end{aligned} \quad (\text{D34})$$

where we used Eq. (D32) and the fact that Q must be a symmetric matrix. The function \mathcal{L}_u is continuous and differentiable with respect to u . Also, duality theory of convex optimization promises that \mathcal{L}_u is concave in u . Hence, in the scenario where the optimal u is not attained for the values $u = \{0, 1\}$, we will have that \mathcal{L}_u is optimized at its stationary point $d\mathcal{L}_u/du = 0$. The optimality condition of our HCRB is hence reduced to finding the roots of the stationary points of \mathcal{L}_u . This concludes our proof of theorem 3 in the main body of the text for any two-parameter estimation problem.

Appendix E: Complexity analysis for Lagrange minimisation

For fixed u , our method to bound the HCRB requires minimising the Lagrangian. We found that this amounts to solving

the Sylvester equation $A^\dagger = -(uY\rho + (1-u)\rho Y)$ for Y . For large dimensional systems, evaluating the HCRB with concrete analytical results becomes increasingly cumbersome. In this scenario, a numerical approach can be used to handle the state diagonalisation to evaluate the HCRB. Here, we bound the complexity of our formalism to evaluating the HCRB numerically.

For large dimensional systems, the Sylvester equation is more efficiently solved by first vectorising the equation to

$$\text{vec}(A^\dagger) = -(\mathbb{1}_D \otimes (1-u)\rho + u\rho^\top \otimes \mathbb{1}_D) \text{vec}(Y), \quad (\text{E1})$$

and then solved using any system of linear equations solver. The Bartels-Stewart algorithm is an efficient and robust numerical solver for the Sylvester matrix for large- D [67, 90], which outperforms well-known primitive implementations of Gaussian elimination. The complexity of the Bartels-Stewart algorithm scales as $O(D^3)$.

To circumvent any time complexity involved, we must have access to the basis that diagonalises the state. In this case, we solved the Sylvester equation analytically. For large D , there are efficient numerical methods that can attain the spectral decomposition of the state in sub-cubic time. For example, with $\rho(\theta) \in \mathbb{H}_D$, the matrix inversion operation can be practically achieved using the Coppersmith-Winograd algorithm, which scales as $O(D^{2.376})$ [91].

As a point of comparison, semi-definite programming (SDP) provide an alternative method to optimisation tasks. SDP programs can be applied to general problems and admit polynomial-time solvers, which highlight the power of this approach. The HCRB was recast as an SDP program in Ref. [44]. For a consistent complexity comparison with our method, we consider the non-trivial case where the state is full-rank. Hence, by observation of Eq. (11) in Ref. [44], the variable X that is optimised has order D^2 terms. Further, notice that for each iteration of the SDP algorithm, the first constraint requires knowledge of the spectral decomposition of a matrix parameterised in terms of X . Therefore, it is easy to observe that this brute force SDP approach has a time complexity greater than $O(D^{2 \times 2.376})$. This indicates at least a quadratic speedup, which amounts to a significant improvement with increasing D .

-
- [1] S. L. Braunstein and C. M. Caves, “Statistical distance and the geometry of quantum states,” *Phys. Rev. Lett.*, vol. 72, no. 22, pp. 3439–3443, 1994.
 - [2] V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology,” *Nat. Photon.*, vol. 5, pp. 222–229, April 2011.
 - [3] G. Tóth and I. Apellaniz, “Quantum metrology from a quantum information science perspective,” *J. Phys. A: Math. Theoret.*,

vol. 47, p. 424006, Oct. 2014.

- [4] C. Helstrom, “The minimum variance of estimates in quantum signal detection,” *IEEE Transactions on Information Theory*, vol. 14, pp. 234–242, March 1968.
- [5] C. W. Helstrom, *Quantum Detection and Estimation Theory*. Academic Press Inc., 1976.
- [6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum*

- Theory*. Springer, 1 ed., 2011.
- [7] K. Matsumoto, “A new approach to the cramer-rao-type bound of the pure-state model,” *J. Phys. A: Math. Gen.*, vol. 35, p. 3111, March 2002.
 - [8] M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference*. World Scientific Publishing, February 2008.
 - [9] R. D. Gill and M. Guta, “On asymptotic quantum statistical inference,” *arXiv.org*, January 2012.
 - [10] D. Burgarth and A. Ajoy, “Evolution-free hamiltonian parameter estimation through zeeman markers,” *Phys. Rev. Lett.*, vol. 119, p. 030402, July 2017.
 - [11] M. Tsang, H. M. Wiseman, and C. M. Caves, “Fundamental quantum limit to waveform estimation,” *Phys. Rev. Lett.*, vol. 106, p. 090401, March 2011.
 - [12] S. Z. Ang, G. I. Harris, W. P. Bowen, and M. Tsang, “Optomechanical parameter estimation,” *New J. Phys.*, vol. 15, p. 103028, October 2013.
 - [13] T. Baumgratz and A. Datta, “Quantum enhanced estimation of a multidimensional field,” *Phys. Rev. Lett.*, vol. 116, p. 030801, January 2016.
 - [14] M. Tsang, “Quantum imaging beyond the diffraction limit by optical centroid measurements,” *Phys. Rev. Lett.*, vol. 102, p. 253601, June 2009.
 - [15] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, “Quantum enhanced multiple phase estimation,” *Phys. Rev. Lett.*, vol. 111, p. 070403, August 2013.
 - [16] C. N. Gagatsos, D. Branford, and A. Datta, “Gaussian systems for quantum-enhanced multiple phase estimation,” *Phys. Rev. A*, vol. 94, p. 042342, October 2016.
 - [17] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, “Optimal measurements for simultaneous quantum estimation of multiple phases,” *Phys. Rev. Lett.*, vol. 119, p. 130504, September 2017.
 - [18] J. Rubio and J. Dunningham, “Bayesian multiparameter quantum metrology with limited data,” *Phys. Rev. A*, vol. 101, p. 032114, March 2020.
 - [19] T. J. Proctor, P. A. Knott, and J. A. Dunningham, “Networked quantum sensing,” unpublished, February 2017.
 - [20] J. S. Sidhu and P. Kok, “Quantum metrology of spatial deformation using arrays of classical and quantum light emitters,” *Phys. Rev. A*, vol. 95, p. 063829, June 2017.
 - [21] J. S. Sidhu and P. Kok, “Quantum fisher information for general spatial deformations of quantum emitters,” *ArXiv*, February 2018.
 - [22] J. R. Jiménez, P. Knott, T. Proctor, and J. A. Dunningham, “Quantum sensing networks for the estimation of linear functions,” *J. Phys. A: Math. Theoret.*, June 2020.
 - [23] M. G. A. Paris, “Quantum estimation for quantum technologies,” *Int. J. Quantum Inf.*, vol. 7, pp. 125–137, 2009.
 - [24] D. A. S. Fraser, “On local unbiased estimation,” *J. R. Stat. Soc. Ser. B*, vol. 26, pp. 46–51, October 1964.
 - [25] M. J. W. Hall and H. M. Wiseman, “Does nonlinear metrology offer improved resolution? answers from quantum information theory,” *Phys. Rev. X*, vol. 2, p. 041006, October 2012.
 - [26] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, “Quantum Limits in Optical Interferometry,” *Progress in Optics*, vol. 60, pp. 345–435, 2015.
 - [27] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1)*. United States: Prentice Hall, 1 ed., 1993.
 - [28] L. Pezzè and A. Smerzi, “Quantum theory of phase estimation,” *Proc. Int. School Phys. Enrico Fermi*, vol. Course 188, p. 691, November 2014.
 - [29] C. L. Degen, F. Reinhard, and P. Cappellaro, “Quantum sensing,” *Rev. Mod. Phys.*, vol. 89, p. 035002, July 2017.
 - [30] J. S. Sidhu and P. Kok, “Geometric perspective on quantum parameter estimation,” *AVS Quantum Science*, vol. 2, p. 014701, February 2020.
 - [31] H. Zhu, “Information complementarity: A new paradigm for decoding quantum incompatibility,” *Scientific Reports*, vol. 5, pp. 14317–14329, September 2015.
 - [32] T. Heinosaari, T. Miyadera, and M. Ziman, “An invitation to quantum incompatibility,” *Journal of Physics A: Mathematical and Theoretical*, vol. 49, p. 123001, February 2016.
 - [33] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, “Compatibility in multiparameter quantum metrology,” *Phys. Rev. A*, vol. 94, pp. 052108–11, Nov. 2016.
 - [34] M. Guță and J. Kahn, “Local asymptotic normality for qubit states,” *Physical Review A*, vol. 73, May 2006.
 - [35] M. Hayashi and K. Matsumoto, “Asymptotic performance of optimal state estimation in qubit system,” *J. Math. Phys.*, vol. 49, no. 10, p. 102101, 2008.
 - [36] K. Yamagata, A. Fujiwara, and R. D. Gill, “Quantum local asymptotic normality based on a new quantum likelihood ratio,” *Ann. Statist.*, vol. 41, pp. 2197–2217, October 2013.
 - [37] Y. Yang, G. Chiribella, and M. Hayashi, “Attaining the ultimate precision limit in quantum state estimation,” *Commun. Math. Phys.*, vol. 368, pp. 223–293, May 2019.
 - [38] J. Suzuki, “Explicit formula for the holevo bound for two-parameter qubit-state estimation problem,” *J. Math. Phys.*, vol. 57, pp. 042201–29, Apr. 2016.
 - [39] M. Bradshaw, S. M. Assad, and P. K. Lam, “A tight cramer-rao bound for joint parameter estimation with a pure two-mode squeezed probe,” *Physics Letters A*, vol. 381, no. 32, pp. 2598 – 2607, 2017.
 - [40] P. A. Knott, T. J. Proctor, A. J. Hayes, J. F. Ralph, P. Kok, and J. A. Dunningham, “Local versus global strategies in multiparameter estimation,” *Phys. Rev. A*, vol. 94, pp. 062312–7, Dec. 2016.
 - [41] J. Řeháček, Z. Hradil, B. Stoklasa, M. Paur, J. Grover, A. Krzic, and L. L. Sánchez-Soto, “Multiparameter Quantum Metrology of Incoherent Point Sources: Towards Realistic Superresolution,” *arXiv.org*, p. arXiv:1709.07705, Sept. 2017.
 - [42] L. Zhang and K. W. C. Chan, “Quantum multiparameter estimation with generalized balanced multimode NOON-like states,” *Phys. Rev. A*, vol. 95, pp. 032321–6, Mar. 2017.
 - [43] T. J. Proctor, P. A. Knott, and J. A. Dunningham, “Multiparameter estimation in networked quantum sensors,” *Phys. Rev. Lett.*, vol. 120, p. 080501, February 2018.
 - [44] F. Albarelli, J. F. Friel, and A. Datta, “Evaluating the holevo cramer-rao bound for multiparameter quantum metrology,” *Phys. Rev. Lett.*, vol. 123, no. 20, p. 200503, 2019.
 - [45] F. Fröwis, R. Schmied, and N. Gisin, “Tighter quantum uncertainty relations following from a general probabilistic bound,” *Phys. Rev. A*, vol. 92, p. 012102, July 2015.
 - [46] R. D. Gill and S. Massar, “State estimation for large ensembles,” *Phys. Rev. A*, vol. 61, p. 042312, March 2000.
 - [47] O. E. Barndorff-Nielsen and R. D. Gill, “Fisher information in quantum statistics,” *J. Phys. A: Math. Gen.*, vol. 33, pp. 4481–4490, June 2000.
 - [48] E. Bagan, M. Baig, R. Muñoz Tapia, and A. Rodriguez, “Collective versus local measurements in a qubit mixed-state estimation,” *Phys. Rev. A*, vol. 69, p. 010304, January 2004.
 - [49] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz Tapia, “Optimal full estimation of qubit mixed states,” *Phys. Rev. A*, vol. 73, p. 032301, March 2006.
 - [50] J. Kahn and M. Guta, “Local Asymptotic Normality for Fi-

- nite Dimensional Quantum Systems,” *Commun. Math. Phys.*, vol. 289, pp. 597–652, Mar. 2009.
- [51] C. W. Helstrom, “Minimum mean-squared error of estimates in quantum statistics,” *Phys. Lett. A*, vol. 25, no. 2, pp. 101–102, 1967.
- [52] H. P. Yuen and M. Lax, “Multiple-parameter quantum estimation and measurement of nonselfadjoint observables,” *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 740–750, 1973.
- [53] V. P. Belavkin, “Generalized uncertainty relations and efficient measurements in quantum systems,” *Theoret. Math. Phys.*, vol. 26, pp. 213–222, March 1976.
- [54] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*. Graduate Texts in Mathematics, Springer, 2013.
- [55] H. Nagaoka, “A new approach to Cramér-Rao bounds for quantum state estimation,” it 89-42, IEICE Technical Report, 1989. reprinted in [8], Chapter 8, pages 100–112.
- [56] M. Hayashi and K. Matsumoto, “Statistical model with measurement degree of freedom and quantum physics,” *Surikaiseki Kenkyusho Kokyuroku*, vol. 96, no. 1055, 1998. reprinted in [8], Chapter 13, pages 162–169.
- [57] M. Keyl and R. F. Werner, “Estimating the spectrum of a density operator,” *Phys. Rev. A*, vol. 64, p. 052311, October 2001.
- [58] M. Hayashi and K. Matsumoto, “Asymptotic performance of optimal state estimation in quantum two level system.” unpublished, August 2006.
- [59] M. G. Genoni, M. G. A. Paris, G. Adesso, H. Nha, P. L. Knight, and M. S. Kim, “Optimal estimation of joint parameters in phase space,” *Phys. Rev. A*, vol. 87, p. 012107, January 2013.
- [60] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*. North Holland, Amsterdam, 1982.
- [61] A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, “On quantumness in multi-parameter quantum estimation,” *J. Stat. Mech.-Theory E*, vol. 2019, p. 094010, September 2019.
- [62] M. Tsang, “The holevo cramér-rao bound is at most thrice the helstrom version.” unpublished, November 2019.
- [63] F. Albarelli and A. Datta, “Upper bounds to the holevo cramér-rao bound for multiparameter quantum metrology.” unpublished, November 2019.
- [64] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, “Compatibility in multiparameter quantum metrology,” *Phys. Rev. A*, vol. 94, pp. 052108–11, Nov. 2016.
- [65] J. Suzuki, “Information geometrical characterization of quantum statistical models in quantum estimation theory,” *Entropy*, vol. 21, p. 703, July 2019.
- [66] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [67] R. H. Bartels and G. W. Stewart, “Algorithm 432: Solution of the matrix equation $ax+xb=c$,” *Commun. ACM*, vol. 15, pp. 820–826, September 1972.
- [68] J. Gorski, F. Pfeuffer, and K. Klamroth, “Biconvex sets and optimization with biconvex functions: a survey and extensions,” *Math. Methods Oper. Res.*, vol. 66, pp. 373–407, June 2007.
- [69] L. Razzoli, L. Ghirardi, I. Siloi, P. Bordone, and M. G. A. Paris, “Lattice quantum magnetometry,” *Phys. Rev. A*, vol. 99, p. 062330, June 2019.
- [70] Y.-X. Liu, A. Ajoy, and P. Cappellaro, “Nanoscale vector dc magnetometry via ancilla-assisted frequency up-conversion,” *Phys. Rev. Lett.*, vol. 122, p. 100501, March 2019.
- [71] S. Pang and T. A. Brun, “Quantum metrology for a general hamiltonian parameter,” *Phys. Rev. A*, vol. 90, p. 022117, August 2014.
- [72] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs, and J. J. L. Morton, “Magnetic field sensing beyond the standard quantum limit using 10-spin noon states,” *Science*, vol. 324, pp. 1166–1168, May 2009.
- [73] Y. Ouyang, “Permutation-invariant quantum codes,” *Phys. Rev. A*, vol. 90, p. 062317, December 2014.
- [74] Y. Ouyang, N. Shettell, and D. Markham, “Robust quantum metrology with explicit symmetric states.” unpublished, 2019.
- [75] D. Gottesman, A. Kitaev, and J. Preskill, “Encoding a qubit in an oscillator,” *Phys. Rev. A*, vol. 64, p. 012310, Jun 2001.
- [76] K. Noh, S. M. Girvin, and L. Jiang, “Encoding an oscillator into many oscillators,” 2019.
- [77] I. L. Chuang, D. W. Leung, and Y. Yamamoto, “Bosonic quantum codes for amplitude damping,” *Phys. Rev. A*, vol. 56, p. 1114, 1997.
- [78] W. Wasilewski and K. Banaszek, “Protecting an optical qubit against photon loss,” *Phys. Rev. A*, vol. 75, p. 042316, Apr 2007.
- [79] M. Bergmann and P. van Loock, “Quantum error correction against photon loss using NOON states,” *Phys. Rev. A*, vol. 94, p. 012311, Jul 2016.
- [80] Y. Ouyang and R. Chao, “Permutation-invariant constant-excitation quantum codes for amplitude damping,” *IEEE Transactions on Information Theory*, vol. 66, no. 5, pp. 2921–2933, 2019.
- [81] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, “New class of quantum error-correcting codes for a bosonic mode,” *Phys. Rev. X*, vol. 6, p. 031006, Jul 2016.
- [82] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, “Quantum computing with rotation-symmetric bosonic codes,” *Phys. Rev. X*, vol. 10, p. 011058, Mar 2020.
- [83] B. M. Terhal, “Quantum error correction for quantum memories,” *Rev. Mod. Phys.*, vol. 87, pp. 307–346, April 2015.
- [84] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, “New class of quantum error-correcting codes for a bosonic mode,” *Phys. Rev. X*, vol. 6, p. 031006, July 2016.
- [85] B. Terhal, J. Conrad, and C. Vuillot, “Towards scalable bosonic quantum error correction,” *Quant. Sci. Tech.*, June 2020.
- [86] Y. Ouyang and N. Rengaswamy, “Weight distribution of classical codes influences robust quantum metrology,” *arXiv preprint arXiv:2007.02859*, 2020.
- [87] E. Deadman and S. D. Relton, “Taylor’s theorem for matrix functions with applications to condition number estimation,” *Linear Algebra and its Applications*, vol. 504, pp. 354 – 371, 2016.
- [88] Y. Ouyang, “Quantum storage in quantum ferromagnets,” *arXiv preprint arXiv:1904.01458*, 2019.
- [89] M. J. Best, *Portfolio Optimization*. Chapman and Hall/CRC, 1 ed., March 2010.
- [90] S. N. G. Golub and C. V. Loan, “A hessenberg-schur method for the problem $ax+xb=c$,” *IEEE Trans. Autom. Control*, vol. 24, pp. 909–913, December 1979.
- [91] D. Coppersmith and S. Winograd, “Matrix multiplication via arithmetic progressions,” *J. Symb. Comput.*, vol. 9, pp. 251–280, March 1990.
- [92] The parameters G and n here are reminiscent of GNU-states in permutation-invariant quantum codes [73]
- [93] A constraint function is $f(x)$ written in the canonical form $f(x) = 0$ or $f(x) \leq 0$